DOCUMENT RESUME

ED 333 039 TM 016 565

AUTHOR Hansen, Joe B.; Hathaway, Walter E.
TITLE A Survey of More Authentic Assessment Practices.
PUB DATE 4 Apr 91
NOTE 92p.; Paper presented at the Annual Meeting of the National Council on Measurement in Education (Chicago, IL, April 4-6, 1991).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150) -- Tests/Evaluation Instruments (160)

EDRS PRICE MF01/PC04 Plus Postage.
DESCRIPTORS *Educational Assessment; Educational Policy; Educational Practices; Elementary Secondary Education; *Evaluation Methods; Foreign Countries; *Mail Surveys; *Measurement Techniques; Organizations (Groups); Questionnaires; School Districts; *Test Use; *Whole Language Approach
IDENTIFIERS *Authentic Assessment; Canada; United States

ABSTRACT

Information was gathered to determine the extent to which more authentic education assessments, including whole language assessments, are being implemented, and to identify which organizations conduct authentic education assessments. Materials used by the surveyed organizations were collected and catalogued into a compendium of authentic assessment techniques. A conceptual framework was developed for characterizing the stage of development an organization attained with respect to authentic assessment; this framework ranged from the initial policy decision through making the assessment operational. A questionnaire was mailed to individuals representing 433 educational organizations and businesses in the United States, Canada, and several other countries. Representatives were asked what they were doing to answer the call for more authentic assessment and, more specifically, how they were addressing the call for more whole language assessment. A total of 110 organizations responded positively to either of the two questions administered, a response rate of 25.4%. Of these, 106 respondents indicated that they were attempting more authentic assessment, with 75 attempting whole language assessment. Forty-one organizations submitted materials to be catalogued. In general, assessments were in the developmental stage, with relatively little full implementation. Seven tables contain study data. Appendix A presents the survey questionnaire, and Appendix B lists 55 abstracts of the materials submitted. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED333039

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Joe B. Hansen

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

A SURVEY OF MORE AUTHENTIC ASSESSMENT PRACTICES

Joe B. Hansen, Ph.D.
Colorado Springs Public Schools

Walter E. Hathaway, Ph.D.
Portland Public Schools

A Presentation in the NCME/NATD symposium:

MORE AUTHENTIC ASSESSMENT: Theory and Practice

Chicago, Ill.

April 4, 1991

TM016565

BEST COPY AVAILABLE

TABLE OF CONTENTS

I. Introduction 1

A. The Need for More Authentic Assessment 1
B. Whole Language Assessment 4
C. The Need for Information on Authentic Assessment Practices 6
D. Purpose of the Study 6
E. The Need for a Conceptual Framework 7
F. The Need to Examine Stages of Development in Authentic Assessment 9

II. Method 11

III. Results 12

A. Scope of Study 12
B. Response Rate 12
C. Question One - Authentic Assessment 13
D. Question Two - Whole Language 15

IV. Discussion 17

V. Conclusions 19

VI. References 21

VII. Appendix-A: The survey 23

VIII. Appendix-B: Abstracts of materials submitted 25

INTRODUCTION

The american education system is under intense scrutiny and pressure from virtually every political, educational, business or interest group in our society, for its alleged shortcomings. Since The publication of A Nation At Risk: The Imperative for School Reform was published in 1983, the challenge to produce students who can compete intellectually in the global society has been the focus and concern of professional educators, politicians and others interested in the preservation of the american standard of living. The call for educational restructuring has grown from a whisper to a tumultuous roar. The burden on the public schools to cure the ills of society has grown in geometric proportions over the past two decades, while the willingness of the public to pay spiraling educational costs continues to diminish. Greater accountability for the educational tax dollar is being demanded by nearly everyone who has a stake in public education at the same that the demand for increased academic achievement is being so emphatically stated. Schools of Choice, voucher systems, site based management, and numerous other "solutions" are being advocated by educational interest groups of all types. Education is becoming the top priority issue for the nation as we approach the twenty-first century.

The Need for More Authentic Assessment

As the pressure mounts for significant educational reforms to occur, the public is demanding more and stronger evidence that such reforms are working to produce students who can think, communicate effectively, solve complex problems and find solutions to the economic, social and political, environmental and other problems facing our society. As all of this has been occurring, educators, over the past decade, have become increasingly aware of the limitations of standardized tests as the means of assessing student performance and evaluating significant educational change. Numerous commissions, study groups and forums have been created to deal with what is perceived to be a major educational problem: the need for greater authenticity in the way we assess student learning and growth. One such commission, the National Commission on Testing and Public Policy (NCTPP) recently concluded a three year study of "...trends , practices and impacts of the use of standardized testing instruments and other forms of assessment in schools, the workplace, and the military." A key finding of this commission was that:

> **"Current testing, predominantly multiple choice in format, is over-relied upon, lacks adequate public accountability, sometimes leads to unfairness in the allocation of opportunities, and too often undermines vital social policies." (NCTPP, 1990, p. ix)**

This broadly based commission included individuals with expertise, interests and experience in a wide variety of fields, including education, business, labor, law, assessment and measurement, and manpower development and training. The NCTPP study concluded that:

> **"To help promote greater development of the talents of all our people, alternative forms of assessment must be developed and more critically judged and used, so that testing and assessment open gates of opportunity rather than close them off." (p. x.)**

The commission identified five key limitations of testing that must be addressed:

1.) Tests are imperfect and therefore potentially misleading as measures of individual performance in education and employment.

2.) Some test uses result in unfair treatment of individuals and groups.

3.) Students are subjected to too much testing in the nation's schools.

4.) Some testing practices in both education and employment undermine important social policies and institutions intended to develop or utilize human talent.

5.) Tests have become instruments of public policy without sufficient public accountability.

Whether one agrees with these findings and recommendations or not, they serve to articulate clearly the frustration of educators and education watchers with the limitations of conventional testing. They also serve to underscore the need for increased relevancy and authenticity in educational assessment tools and techniques.

In recent years, numerous authors have reviewed the deficiencies of standardized tests and called for greater authenticity in educational assessments (Archibald and Newmann, 1988, Cannell, 1989; Shepard, 1989; Neil and Medina, 1989; Wiggins, 1989, 1990.) These authors and numerous others have called for the development and use of assessment techniques that are "more authentic." This means that the assessments used should have certain characteristics that simulate the conditions a student would experience in applying his/her knowledge or skill in a real world environment. Thus the student has an opportunity to demonstrate his/her knowledge, extend it and apply it. Authentic assessments, as defined by Archibald and Newmann (1988) must meet three criteria: 1.) production of discourses, things or performances, 2.) flexible use of time, and 3.) collaboration. These criteria are based on the concept of "disciplined inquiry" in the real world environment, outside of the

classroom. In describing the Connecticut Common Core of Learning's "enriched performance tasks", Baron (1990), lists features such as: grounded in real world contexts; involve sustained work; based upon the most essential aspects of the discipline(s) being assessed; are broad in scope; blend essential content with essential practice; present non-routine, open-ended and sometimes loosely structured problems that require the student to both define the problem and determine a strategy for solving it; encourage group discussion and "brainstorming"; require students to make, explain, and defend their assumptions, predictions and estimates; stimulate students to make connections and generalizations that will increase their understanding of important concepts and processes; are accompanied by explicitly stated scoring criteria related to content, process, group skills, communication skills, and a variety of motivational dispositions and "habits of mind"; spur students to monitor themselves and to think about their progress; necessitate that students use a variety of skills both for acquiring information and for communicating their strategies, data conclusions, and reflections.

> "Authentic evaluation of educational achievement directly measures actual performance in the subject area. Standardized multiple choice tests, on the other hand, measure test taking skills directly, and everything else either indirectly or not at all." (FairTest, undated)

An authentic assessment task should be worthwhile, significant and meaningful, and should provide substantive information (Archibald and Newmann, 1988).

Authentic assessments require students to demonstrate what they know and can do rather than to select a "correct" answer from a list of alternatives. In this sense they are performance based. They have diverse and varied formats that are derived from and relevant to the skill or knowledge being tested. For example, a direct writing assessment that evaluates a student's performance on a sample of the student's writing is an authentic assessment. They involve complex and sustained tasks that tend to call for higher order cognitive skills, than recall, recognition, or simple deductive reasoning. Some higher order cognitive skills described in the book Dimensions of Thinking: a framework for thinking and instruction (Marzano, Brandt, Hughes, Jones, Pressiesen, Rankin, Suhor, 1988) that are associated with authentic assessments include: information gathering, generating, analyzing, integrating, and evaluating.

More recently, numerous organizations have emerged to join the battle to reduce our dependency on standardized testing and increase the authenticity of assessments, including the National Education Association (NEA), The National Center for Fair and Open Testing (FairTest), The National Urban Alliance (NUA) and the American Federation of Teachers (AFT).

The response to this overwhelming ground swell of concern regarding the need for better assessment tools has been widespread. The National Assessment for Educational Progress (NAEP), operated by Educational Testing Service (ETS), the largest test publisher in the United States, has responded by incorporating open ended performance type items in the 1990 math and language assessments, with plans to expand the number of these items in the 1992 assessment. Other test publishers are following suit by offering commercial packages incorporating features of "authentic assessment." Additionally, numerous state assessment programs are incorporating more authentic techniques.

Whole Language Assessment

The trend toward "authentic assessment" has been parallelled by a call for greater authenticity in instruction, in an effort to have instruction correspond more closely to emerging knowledge on human learning. The recognition of multiple forms of intelligence (Gardner, 1983) has had a pronounced effect on pedagogical thought in recent years, resulting in efforts to engage students in a variety of modes of instruction in an effort to develop the whole child and provide the student an opportunity to exercise his or her strengths in the mastery of the instructional material experienced. Performance based assessments have both led and followed from the development of performance based approaches to learning.

There is perhaps no discipline more fundamental to the development of the human intellect than reading. Without adequate reading skill, most knowledge remains locked away out of reach of those who would seek it. Few would argue that even in the electronic age, reading is not the most important of skills a student must acquire early in his/her schooling, if that student is to succeed in school, and indeed, in life. Therefore the teaching of reading has been a central concern of education from the beginning of the written word.

The "whole language" approach to teaching reading is based on the precept that reading, writing and speaking are integrally linked to one another and to learn each is to learn something of the others. Whole language is a "constructivist" approach which advocates that understanding can best be accomplished through the act of reading itself rather than through the development of discrete component skills, such as word recognition, phonics, spelling and so forth. This approach relies on the use of contextual cues, prior knowledge and potential helpers for the construction of meaning from written material (Valencia and Pearson, 1987.) It calls for the integration of reading throughout the curriculum in social studies, science, mathematics and other subject areas as a fundamental way of learning. It is as much a theory of learning as a theory of language development. (Harste, 1989). Whole language is said to integrate the holistic psychological research

of Piaget, Vygotsky and schema theorists with the social, functional-linguistic research of Halliday. It goes beyond the positive, child-centered education movements of the past to integrate scientific concepts and theories of language processes, learning and cognitive development, teaching and curriculum into a practical philosophy to guide classroom decision making (Goodman, K., 1989). So whole language is much more than an instructional technique for language development, it is a theory, philosophy and a major educational movement that is gradually but with certainty, replacing traditional, basal reader, skills based approaches to reading and language instruction.

This approach to instruction poses a challenge to educators to find appropriate assessment techniques for measuring student progress in reading and language development. It is antithetical to the concept of whole language instruction to use multiple choice tests to assess reading growth, since to do so would require the specification of sub-skills within the overall level of language skill development. Whole language assessment is based on "kid watching" (Goodman, K., 1989). This type of evaluation is a continuous ongoing, integral process in which teachers are involved in evaluating all aspects of the curriculum. It is based on a "double agenda" wherein students are learning through teacher-student interaction and there is continuous reflection on and questioning of the instructional process by both teacher and student. Teachers monitor student performance and ask such questions as "who is getting things done? How are students concepts and hypotheses changing? Who seems confused? How did things go in our discussion group? Did I organize the writing area so those who wanted to write could do so? " and so on (Goodman, Y., 1989.) The key techniques used by teachers for evaluating student growth and their own efficacy are "observing, interacting and analyzing."

Teachers using the whole language approach have been on their own to devise and experiment with appropriate assessment tools and procedures during the course of instruction. Miscue analysis, portfolios, informal and formal observation techniques and other non-standardized approaches have been the standby assessment tools of the whole language teacher. While having face validity, these techniques are criticized for their subjectivity, lack of reliability and failure to yield aggregatable data. The Northwest Evaluation Association (NWEA) is one organization that has made a commitment to addressing some of the issues associated with whole language assessment. The NWEA, has, over the past year, held a series of working retreats, involving educational practitioners, evaluators and researchers in finding solutions to the problems of

aggregating portfolio data. The results of this work to date, look very promising.[1]

Whole language assessment, therefore is a special case of "authentic assessment". It is an area of intense need, for which few effective tools have been well developed. Some research and development has taken place in a few of the state operated large scale assessment programs such as those in Michigan, Illinois and California, but relatively few tools are available for use at district or school level. Direct writing assessment is the formal measurement technique most often associated with whole language on a large scale, although it is not used exclusively for whole language assessment.

The Need for Information on Authentic Assessment Practices

We seem to have reached the point where it is virtually impossible to find anyone in or out of education who is willing to defend the practice of using standardized tests, regardless of the circumstances. Perhaps it is time to slow down and take stock of where this movement is leading and assess the extent to which public education will be likely to benefit from this rapidly growing trend in assessment. A starting point for ascertaining where we are headed might be to first establish just where we are presently. This is why, in the early Spring of 1990, Dr. Walter Hathaway, Director of Research, Evaluation and Testing for the Portland Public Schools, decided to conduct a broadly based survey of the state of the art in authentic assessment. I contracted with Dr. Hathaway and Portland Public Schools to conduct any needed follow up, analyze and compile the results and write the study report.

Purpose of the Study

The survey had two major objectives. The first was to gather information that would enable us to know the extent to which more authentic assessment techniques, including whole language assessments, were being implemented and by what organizations. Our second objective was to elicit materials from organizations engaged in authentic assessment efforts, that could be catalogued into a compendium of authentic assessment tools and techniques. We felt that such a compendium could become a resource base for educators interested in developing or studying authentic assessment. A simple questionnaire, requiring minimal effort and respondent time was designed to procure this information.

---

[1] For further information contact Allan Olson, Executive Director, NWEA, 5 Centerpointe Drive, Lake Oswego, OR 97035.

The Need for a Conceptual Framework

As researchers we felt a need for a conceptual framework (Miles and Huberman, 1985) to guide us in classifying and interpreting the responses we would obtain from the survey and to provide a schemata for classifying the resource documents we would collect. Therefore we explored the authentic assessment literature in search of ideas for such a framework.

Basic precepts underlying a conceptual framework: Any attempt at classifying authentic assessment efforts must begin with the question of purpose. What is the purpose of the assessment in question?

Several purposes for assessment have been identified and stand as conventions. One taxonomy of purpose can be adapted from Stiggins and Bridgeford (1985) based on research on classroom assessments. These researchers defined classroom testing in terms of five purposes based on the types of decisions that teachers make: diagnosing, grouping, grading, evaluating, and reporting. Distinctions are also made across types of tests, e.g. publishers', teacher made, performance, etc. This has been extended as follows: Diagnosing strengths and weaknesses of individual pupils, diagnosing group needs, grouping students for instruction, determining the achievement potential of students, assigning grades to students, evaluating the instructional unit to see if it worked, communicating academic expectations, and controlling and motivating students.

Other purpose factors identified by Stiggins (1987), affecting assessment, include identifying students for special services, communicative affective or behavioral expectations, and providing test taking experiences.

Yet another taxonomy of purpose is offered by Airasian (1984) in which two general levels of assessment are offered, 1.) externally mandated assessments, and 2.) teacher used assessments. The former includes high school graduation tests, preschool screening procedures, and test to allocate resources, such as for a remedial reading program. These are typically mandated by an authority external to the school, typically the state school board or the legislature. To this list one could add statewide assessment programs and tests required by state accountability programs. These assessments serve the purpose of providing policy makers, data for monitoring the progress of schools and districts on state mandated goals or curriculum. Airasian's second level, those assessments used by teachers in their classrooms, are the assessments used to guide their instruction, maintain order, and to assign grades in their classrooms. These two levels of assessment have distinctively different characteristics.

Externally mandated assessments involve tests, regulations, objectives and passing scores set by the external authority. Their

purpose is often external and may include a moral or managerial dimension. They rely heavily on a single test score and a single cutoff point for decision making.

Classroom assessments are performed on a day to day basis. They vary in form, e.g. standardized, non-standardized, teacher made (oral, observation, paper-pencil, essay or objective), performance checklists, rating forms, work samples, anecdotal records.

As the work of Stiggins and Airasian emphasize, the purpose of assessment is a key determinant in deciding what type of assessment to employ in a particular situation. Obviously, purposes vary across a broad spectrum and carry implications based on whether they are *internally* focused or *externally* imposed.

Another dimension of assessment to be considered in developing a conceptual framework is the type of response required. Much of the criticism directed toward conventional assessment techniques is based on the belief that multiple choice, selected response formats limit the cognitive level of a student's response to that of simple recall, or at best, lower forms of reasoning associated with eliminating the least plausible answers. Selected response formats allegedly don't allow the student to express what he/she knows about the subject in his/her own way. Therefore a useful dimension for classifying authentic assessment efforts, should be the extent to which they call for constructed rather than selected responses. This is not to say that all tests that call for selected responses are inherently "non-authentic." We will save the debate on that issue for another time. It is true however, that authentic assessments are generally perceived to be based on constructed rather than selected responses. Therefore assessments purporting to be more authentic but not requiring constructed responses would have to be subject to special circumstances.

A third property of authentic assessments could be considered to be the level at which they were intended to be interpreted, e.g. individual pupil, classroom, school, district, state, or nation. Portfolios, for example, might be very useful at individual pupil level, useful at class level, perhaps somewhat less useful at school level and present ever increasing difficulties for interpretation as higher levels of aggregation are desired. Some notable exceptions have been reported and experimentation is under way in several places such as NAEP, Vermont State Assessment, and California State Assessment programs to incorporate open-ended responses formats, portfolios and other non-conventional techniques in data to be aggregated at state levels and higher.

No conceptual framework for classifying responses to the call for authentic assessment should be considered unless it takes into account the subject matter or content area being assessed. It is obvious that some subjects lend themselves more readily to more authentic approaches than others do. Direct writing assessment,

for example, has become the standard for assessing student writing ability.

These considerations led us to develop and discard, through an iterative process of trial and error several conceptual frameworks before arriving at one that had the practical utility we sought. The structure we arrived at is illustrated in the chart shown in Table I.

This framework contains five dimensions and could be represented graphically as a pentagon. Each dimension is somewhat independent of the others, although they frequently interact. In classifying the sample tests and descriptive documents, only the first two levels of each dimension were used. This was done because of the lack of uniformity in the amount of detailed information in the descriptions of the assessment tools provided by the various submitters.

The Need to Examine Stages of Development in Authentic Assessment

Hall and Loucks (1977) created a model for tracking an innovation through several stages as it is implemented. This model, the Concerns Based Adoption Model (CBAM) allows the researcher to document the progress of the innovation in terms of the concerns and behaviors of the people who are attempting to implement it. While the CBAM offers an excellent means of tracking the progress of an innovation through its levels of implementation and use, it requires intensive on-site interviews with those persons most affected by the innovation. For our purposes, we needed a way of characterizing the level of implementation without conducting the on-site research required to document the process. For this reason we conceived a hypothetical process of implementation based on a hypothesized cycle of development that flows through five pre-operational stages before the assessment tool or technique reaches a fully operational stage. The process is based on a set of assumptions that result in a linear flow model of developmental stages, somewhat analogous to the CBAM.[2] The assumptions are as follows:

1. Prior to the adoption or implementation of authentic assessment, one of two conditions must exist: either a **policy** decision must be made, to begin to explore the potential of

---

[2] This model, tentatively called Stages of Development (SoD), is itself, still being developed and will be discussed in greater detail in later writings. It is employed here as a heuristic device, to provide a means of characterizing the extent to which an organization has implemented authentic assessment tools and techniques. We make no claims regarding its scientific validity at this time.

authentic assessment, or some **support base** for this approach must exist within the organization.

2. Once the policy or support base is in existence, **research**, either formal or informal may occur in an overt fashion.

3. The research will likely lead to a decision to either begin implementation or seek a new approach. If implementation is chosen then, **training** will need to be provided on how to implement the authentic assessment technique. The training could either be preceded or followed by a **planning** process for implementation.

4. **Development** of the assessment technique may occur concurrent with the planning/teacher training stage or follow from it. Obviously this must precede implementation. Development, in this case could include the purchase or acquisition of a commercial product.

5. **Implementation or pilot testing** may begin once the development has occurred.

6. The authentic assessment technique may become **operational** only after the previous five stages have occurred in some order or fashion.

This set of assumptions provides a heuristic model for characterizing the stage of development an organization has attained with respect to authentic assessment. We argue that the authentic assessment technique cannot become a routine or culturally assimilated practice until these stages have been addressed either directly or vicariously within the organization.

METHOD

A simple questionnaire, consisting of two questions, was mailed out to individuals representing 433 educational organizations and businesses in the United States, Canada and to the education ministries in several foreign countries. Recipients were identified from the mailing lists of the National Association of Test Directors (NATD), Directors of Evaluation (DRE), foreign ministries of education (FME) and the researchers' personal contacts in the measurement and evaluation professional community. The questionnaire, a copy of which is included in Appendix A of this report, asked the following questions:

1. Are you or your organization doing anything to respond to the call for "more authentic educational measurement?"

TABLE I. CONCEPTUAL FRAMEWORK FOR CLASSIFYING "AUTHENTIC ASSESSMENT" EFFORTS

Dimension I. Purpose of Assessment

A. Pupil needs analysis

1.) Content needs - what to teach. (Where is the pupil in the curriculum-skill hierarchy?)

2.) Process needs - how to teach. (Where is the pupil in terms of receptivity to the instructional processes being used, e.g. aural-visual, amount and frequency of practice, etc.)

3.) Cognitive functions, e.g.memory, reasoning, inference, transfer, etc..

4.) Psychomotor needs and level of functionality

B. Selection of pupils for special programs (gifted, remedial, etc.)

C. Grouping for instruction

D. Certification of skills, knowledge, proficiency levels, mastery, etc.

E. Curriculum and instructional program evaluation

F. Policy guidance (district, state, national, etc.)

Dimension II. Scale of Assessment (Level of aggregation and reporting)

A. Individual pupil

B. Classroom or instructional unit

C. Grade level within school

D. School

E. District

F. State

G. National

H. International

Dimension III. Response Type

A. Constructed

1.) Controlled
2.) Open ended
3.) Other

B. Selected

1.) Multi-choice
2.) True-false
3.) Completion
4.) Matching
5.) Other

Dimension IV. Content Area

A. Language Arts

B. Mathematics

C. Reading

D. Science

E. Social science

F. Other

Dimension V. Academic Level

A. Early childhood (pre k - 2)

B. Primary (g1-3)

C. Intermediate (g 4-6)

D. Junior High/Middle (g6-9)

E. High School (g9-12)

F. Post high school

2. Are you or your organization doing anything to respond to the call for "more appropriate assessment of the whole language approach to language arts instruction?"

Respondents were asked to describe the nature of their response if they answered "yes" to either question, and to provide the name, address and phone number of the appropriate contact person in their organization for possible follow up. The initial response level was somewhat disappointing at approximately 20 percent, therefore a limited follow up was planned. At that point we decided that we would focus the follow-up on local school districts, that had not responded to the first mailing. In September, 1990, 125 questionnaires were mailed to a randomly selected sub-sample of the school districts that had not previously responded.

Response data were entered into a PC database so that they could be sorted by respondent group. Assessment materials received from respondents were reviewed and abstracts were developed, summarizing their key characteristics, including information on the submitting organization and a contact person within the organization. The abstracts and materials were then catalogued by the descriptors developed from the conceptual framework to create a library of assessment materials.

RESULTS

Scope of study

The study targeted 433 educational organizations. The number of institutions targeted and response rate by category is shown in Table II. Of the total target group, 187 institutions were public school districts in either the U.S. or Canada, representing 43.2 percent of the target group. Of the remaining groups, 13 percent were colleges or universities. Educational research and development agencies(public and private) comprised 7.2 percent of the total respondent group, while state and canadian provincial education agencies comprised 6.7 percent. Professional associations and certification groups also comprised 6.7 percent of the respondent pool and publishers made up 6.0 percent. A variety of "others" made up the rest. These included intermediate agencies and private schools, councils, commissions and educational forums, smaller private consulting firms, software firms and some other unique designations.

Response Rate

A total of 110 organizations responded positively to one or other of the two questions asked for a 25.4 percent rate of return after follow-up. Of those responding, 106 or 24 percent of the

target population responded positively to question one: "Are you or your organization doing anything to respond to the call for 'more authentic' educational measurement?" Forty one of the 106 or 39 percent, submitted materials for review. Seventy five organizations responded positively to question two: " Are you or anyone in your organization doing anything to respond to the call for 'more appropriate' assessment of whole language approach to language arts instruction?" This represented 17.3 percent of the target group and 68 percent of the total number of respondents to either question. Sixty nine organizations responded positively to both questions, representing 16 percent of the total target group. These data are displayed in Table II.

Table II. Summary of responses to Authentic Assessment Survey by organizational grouping.

Responses to Question One "Authentic Assessment"

TARGET GROUP Organization/institutions receiving questionnaire	Target Group N	Percent of Total	Number "YES"	Percent Target Group	Percent "YES"
Public School Districts (U.S. & Canada)	187	43.19	67	35.83	63.21
Colleges and Universities	56	12.93	9	16.07	8.49
Educational R & D Organizations (public & private)	31	7.16	2	6.45	1.89
State Education Depts. & Canadian Provinces	29	6.70	15	51.72	14.15
Professional Associations - Certification Groups	29	6.70	3	10.34	2.83
Test Publishers	26	6.00	5	19.23	4.72
Councils, Commissions & Forums	9	2.08	1	11.11	.94
Intermediate Education Agencies & Private Schools	7	1.62	2	28.57	1.89
Other	59	13.63	2	1.69	1.89
Total	433	100.00	106	N.A.	100.01*

* Total exceeds 100.00 due to rounding.

Question One - "Authentic Assessment"

Composition of Respondent Group: As shown in Table II 67 or 63 percent of the total number of positive responses to question one were from public school districts. This figure represented 36 percent of the school districts targeted. Seven of 12 Canadian local districts responded positively for a rate of 58 percent. State departments of education and canadian provincial agencies reported higher levels of activity (52 percent) than did any other group identified in the study. Local districts, both american and canadian were second (36 percent.) Out of five canadian provincial agencies reporting, four were positive for an 80 percent rate. Three of 29 professional associations responded positively for a rate of 10 percent. Five of 25 publishers who responded were positive for a 20 percent rate. Only 9 of the 56 respondents representing colleges and universities were positive, for a rate of 16 percent.

Stage of Development: Positive respondents were at many different stages in the development of their authentic assessments, ranging from collecting information to refining operational systems. These data are displayed in Table III. Responses were approximately equally distributed among research (20 percent), development (22 percent), and operational systems (22 percent). Other stages included implementation/pilot (10 percent), developing a policy/support base (2 percent) and not specified (7 percent).

Table III. Stage of Development for Respondents Answering "Yes" on Question One -Authentic Assessment. (Highest level reported)

Stage of Development	Number	Percent
Develop Policy/Support Base	2	.02
Research	21	.20
Planning/Training	19	.18
Development	23	.22
Implementation/Pilot	11	.10
Operational	23	.22
Not Determined	7	.07
Total	106	100

Types of Assessments: Surprisingly, 32 percent of the total number of responses, including multiple statements from the same organization, did not specify the type of assessment they were working with. Direct writing was the most frequently reported type of assessment at 24 percent of the total number of responses recorded. Classroom paper and pencil tests shared the least frequently reported assessment type with norm referenced at 1 percent each. It is interesting to note that norm-referenced assessment was reported as an attempt at more authentic assessment by some respondents. These data are revealed in Table IV.

Grade Levels: As indicated in Table V, fifty five of the 106 positive respondents on question one, or 52 percent did not specify the grade levels at which their authentic assessment efforts were targeted. Twenty seven respondents or 25 percent identified both elementary and secondary grades as their focus. Elementary grades were the focus of 10 percent of the positive responders, while secondary grades were focused on by 8 percent. Only 2 percent were planning to conduct assessments at pre-k or kindergarten levels.

Table IV. Types of Authentic Assessment Reported

Assessment Type	Number	Percent
Standardized(non-specific)	4	.03
Criterion Referenced	5	.04
Norm Referenced	1	.01
Performance Based	20	.15
Direct Writing	33	.24
Portfolio	24	.18
Other	4	.03
Not specified	44	.33
Paper & Pencil(non-spec)	1	.01
Total	135*	101**

* Includes some multiple statements from within same organization.

** Total exceeds 100 due to rounding.

Content Area: Not surprisingly, the most frequently reported area of "authentic assessment" was "language arts", including direct writing assessment. Forty nine of 171 total responses to question one cited language as the content area in which they were working. The second most frequently cited content area was mathematics with 30 responses. These data are shown in Table VI.

Question Two - "Whole Language"

Forty seven of the 187 public school districts targeted responded positively to question two dealing with whole language assessment. As shown in Table VII, this represents 25 percent of the districts targeted and 62.67 percent of the 75 positive responses received. This was by far the highest percentage of positive responses for any group on question two. State departments and canadian provincial ministries of education were next with a 12 percent response. Colleges and universities followed with an 8 percent positive response.

Table V. Responses to Authentic Assessment Question by Grade Level Grouping

Grade Level	Number	Percent
Pre-k - Kindergarten	2	.02
Elementary	11	.10
Secondary	8	.08
Elementary & Secondary	27	.25
Not Specified	55	.52
Not Applicable	3	.03
Total	106	100

Table VI. Responses to Authentic Assessment Question by Content Area

Content Area	Number	Percent
Reading	13	.08
Mathematics	30	.18
Language Arts*	49	.29
Science	18	.11
Social Studies	13	.08
Other	12	.07
Not Specified	36	.21
Total**	171	100

* Language Arts includes writing

** Total exceeds number of "yes" responses due to multiple content areas for some.

Table VII. Summary of responses to Authentic Assessment Survey by organizational grouping.

Responses to Question Two: "Whole Language Assessment"

TARGET GROUP Organization/institutions receiving questionnaire	Number Responding	Percent of Total	Number "YES"	Percent of Group	Percent "YES"
Public School Districts (U.S. & Canada)	187	43.19	47	25.13	62.67
Colleges and Universities	56	12.93	6	10.71	8.00
Educational R & D Organizations (public & private)	31	7.16	3	9.68	4.00
State Education Depts. & Canadian Provinces	29	6.70	9	31.03	12.00
Professional Associations - Certification Groups	29	6.70	0	.00	.00
Test Publishers	26	6.00	5	19.23	6.67
Councils, Commissions & Forums	9	2.08	0	.00	.00
Intermediate Education Agencies & Private Schools	7	1.62	3	42.86	4.00
Other	59	13.63	2	3.39	2.67
Total	433	100.00	75	N.A.	100.00

DISCUSSION

The methodology of the study and the nature of the responses limited the interpretations and conclusions somewhat. Therefore our interpretations and conclusions will be restricted to some of the more obvious outcomes.

Response Rate: Despite the proliferation of rhetoric around the issue of increasing the authenticity of educational assessments, our study found a relatively low percentage of educational institutions and organizations that reported being involved with increasing assessment authenticity in substantive ways. The highest level of activity was found among public school districts in the U.S. and Canada. State Departments and canadian provincial ministries were second. Canadian provinces and districts reported a higher level of activity than did their american counterparts. Whether these results are subject to sampling bias or not is not known, but assuming they are not, we would conclude that the canadians seem to be more involved in improving the authenticity of their educational assessments than we are in the United States.

Unfortunately, although 13 foreign educational ministries were also targeted in the study, only one responded and that response was negative. Therefore comparisons with other countries were not possible.

The educational publishers in our sample reported less activity than either of the public education agency groups. Although all major test publishers in the United States were included in the sample, only five responded positively to the "more authentic assessment" question. Among those not responding were some

extremely large and prominent publishers who have stated publicly and in the media that they support the trend toward more authenticity and have programs under way to improve the authenticity of their tests. This serves to illustrate one of the difficulties of conducting such a survey. Recipients of the survey must feel motivated to respond. Otherwise one cannot infer that the responses are representative of the groups included and the percentage of positive responses must be interpreted as a conservative estimate of the true level of activity. It seems reasonable to assume that those organizations who were actively involved in more authentic assessment would have been more highly motivated to respond those who were less involved. If so, this would explain why the overwhelming majority of respondents answered "yes" to the authentic assessment question. The same assumption would apply to whole the language question.

Only 22 percent of the "authentic assessment" responses were from organizations in the "operational" stage of development. This represented only 5.5 percent of the total target group and 20.8 percent of the responding group. An equal number indicated that they were in the process of developing authentic assessments. Therefore if this survey were to be repeated, one or two years in the future, one would probably find a much higher number of organizations with operational programs of authentic assessment.

Although the "whole language" approach to language arts instruction has been around for several years, a relatively small number of organizations including school districts, publishers and universities reported that they were doing anything to respond to the call for more authentic assessment in this area. Far fewer respondents reported activity in this area than in the more general area of authentic assessment. Only 70 percent as many public school districts reported work in whole language assessment and 60 percent as many state departments and canadian provinces, as did in authentic assessment. This may be attributable to two possible explanations. First, is the possibility that whole language has not been as widely adopted as we initially thought. Perhaps more traditional skills based approaches to reading and language arts instruction still predominate. Second, is the possibility that little attention is being directed toward the assessment issue in whole language instruction, even in those districts and other organizations where it has been adopted. The latter would appear to be consistent with the philosophy of some whole language advocates such as Kenneth Goodman, who has said:

> "When skeptical administrators and colleagues say to such innovative teachers, 'Where is the proof that what you are doing works?' teachers have a right to answer that the proof is in their classrooms and their pupils." (Goodman, K., 1989.)

In the same article as quoted above, Goodman implies that the only appropriate means of assessment for whole language instruction is through learner self reflection and teacher - learner interaction. He seems to imply that formal assessment is incompatible with the theoretical basis of whole language instruction.

CONCLUSIONS

Within the limitations of the survey methodology employed, and the literature examined, we feel justified in drawing the following conclusions.

In general, the state of the art in authentic assessment, at the time of the study, could best be characterized as "developmental." Although respondent organizations were spread across a broad range of stages of development in their authentic assessment efforts, there is more evidence, in the target population, of planning for or contemplating transition to authentic assessment than there is of effective implementation . This study failed to reveal widespread implementation of authentic assessment in the target population, at least in the United States.

Canada seems to be somewhat ahead of the United States in large scale implementation of authentic assessment techniques. Whereas in the United States, the greatest proportion of activity in authentic assessment is found in local districts. In Canada the provincial educational agencies are the locus of most activity. This have more to do with the organizational governance structure of the respective school systems than any other factor.

Efforts to implement authentic assessment of whole language instruction, were less prevalent in our study than were efforts at other types of authentic assessments. This may reflect a lower level of implementation of whole language instruction than was anticipated, or it may be that whole language practitioners have little interest in formalized assessment techniques.

As evidenced in the materials submitted in support of positive responses to the authentic assessment question, there is an extremely wide range of sophistication in the authentic assessment efforts reported. Anomalies appeared in the form of criterion and norm referenced testing materials in the authentic assessment materials either described or submitted by respondents, indicating a lack of understanding of, or familiarity with, accepted criteria for authentic assessments.

There are many good models for authentic assessment in place, especially in the SEAs and canadian provincial agencies. Greater effort should be directed toward monitoring the development of authentic assessment and capturing and sharing information about existing models and techniques to facilitate implementation,

prevent duplication of effort and foster collaboration. Such an effort would save taxpayer dollars, research and development time and broaden the professional knowledge and scientific knowledge base regarding authentic assessment.

Developing a framework for classification of authentic assessment efforts proved difficult. However, the framework we developed should prove useful to other researchers, as a starting point for examining the characteristics and uses of authentic assessments. We hope that others will use this framework, adapting and modifying it as needed.

The hypothetical Stages of Development model was validated in this application, as having heuristic utility for characterizing the stage of development in which an organization is operating, as it moves toward full implementation of authentic assessment. Further work with the SoD model should be conducted to refine it as a research tool.

References

Airasian, P. (1984). Classroom Assessment and Educational Improvement. Keynote address for conference: Classroom Assessment: a key to educational excellence. Portland, OR. Northwest Regional Educational Laboratory.

Archibald, D. and Newmann, F. (1988). Beyond Standardized Testing: Assessing Authentic Academic Achievement In the Secondary School. Reston, VA: National Association of Secondary School Principals.

Berk, R., in Berk, R. Ed. (1986). Performance Assessment: methods and applications. Baltimore. Johns Hopkins University Press.

Cannell, J. (1989). How Public Educators Cheat on Standardized Tests. Albuquerque, NM. Friends for Education.

Gardner, H. (1983). Frames of Mind: The theory of multiple intelligence. New York. Basic Books.

Goodman, K. (1989). Whole-Language Research: Foundations and Development. The Elementary School Journal, 90(2), November, (208-220).

Goodman, Y. (1989). Evaluation of Students: Evaluation of Teachers, in Goodman, K., Goodman, Y. and Hood, W. The Whole Language Evaluation Book. Portsmouth, NH. Heinemann.

Hall, G. and Loucks, S. (1975). A Developmental Model for Determining Whether the Treatment is Actually Implemented. American Educational Research Journal, 14, (3), 263-276.

Harste, J. (1989). Commentary: The Future of Whole Language. The Elementary School Journal, 90(2), November, (243-249).

Miles, M. and Huberman, A. (1985). Qualitative Data Analysis. Beverly Hills. SAGE Publications.

National Commission on Testing and Public Policy (1990). From Gatekeeper to Gateway: transforming testing in America. National Commission on Testing and Public Policy, Boston College, Chestnut Hill, MA.

Neil, M. and Medina, N. (1989). Standardized Testing: Harmful to Educational Health. Phi Delta Kappan, May (688-697).

Shepard, L. (1989). Why We Need Better Assessments. Educational Leadership, 46(7), April, (4-5).

Stiggins, R. and Bridgeford, N. (1985). The Ecology of Classroom Assessment. Journal of Educational Measurement, 22(4), Winter, 271-286.

Stiggins, R. (1987). Profiling Classroom Assessment Environments. Paper presented at annual meeting of National Council on Measurement in Education, Washington, D.C.

Valencia and Pearson (1987). Reading Assessment: Time for a change. The Reading Teacher, April, (726-732).

Wiggins, G. (1989). Teaching to the (Authentic) Test. Educational Leadership, 46(7), April, (41-49).

Wiggins, G. (1990). Reconsidering Standards and Assessment. Education Week, January 24, (36,25).

Appendix A

Survey Questionnaire

QUESTIONNAIRE

1. **Are you or anyone in your organization doing anything to respond to the call for "more authentic educational measurement?"**

Yes ____ No ____

If yes, please briefly describe the nature of the response you are making, below.

_____
_____
_____
_____
_____

If yes, please provide the name, title, address, and phone number of a contact person below.

Name:
Title:
Address:

Phone #

Please attach any relevant documentation.

2. **Are you or anyone in your organization doing anything to respond to the call for "more appropriate assessment of the whole language approach to language arts instruction."**

Yes ____ No ____

If yes, provide briefly describe the nature of hte response you are making, below.

_____
_____
_____
_____
_____
_____

If yes, provide the name, title, address, and phone number of a contact person below.

Name:
Title:
Address:

Phone #

Please attach any relevant documentation.

Appendix B

Abstracts of Response Descriptions and Materials Submitted, Alphabetically by Organization

ORGANIZATION: Alberta Canada Department of Education

CONTACT PERSON: Frank G. Horvath, Director, Student Evaluation
Devonian Building, West Tower
11160 Jasper Ave.
Edmonton, Alberta, CANADA T4K0L2

PHONE: (403) 427-2948

ABSTRACT: The Alberta Department of Education is attempting to incorporate authentic measurement and whole language approach in the tests, diagnostic materials, and examinations they develop with the assistance of classroom teachers. The student evaluation branch is responsible for three related programs: the achievement testing program, the diagnostic evaluation program, and the diploma evaluations program. Sample tests, examinations and supporting documents for each of these three programs were submitted. The achievement testing program is designed to help elementary and junior high school teachers and administrators see how well their students are doing compared with the rest of the province. Achievement tests and language arts, social studies, mathematics and science are administered on a rotating basis at grades 3, 6, and 9. Within the constraints of paper and pencil testing, Alberta has found ways of incorporating and relating three of the language strands, reading, writing and viewing. Teacher training focuses on student strengths, not their weaknesses. Reading passages selected for both parts of the language arts achievement tests are examples of authentic literature. Although multiple choice questions are used, they have been carefully designed to tap increasingly higher level thinking skills. Performance based assessments in math and science are being field tested. Alberta is also developing a strategy for assessing portfolios of students' work at grade three. A diagnostic evaluation program assists teachers to assess strengths and weaknesses of individual students and to provide appropriate instructional activities to help students build upon their strengths and improve on their weaknesses. Two other diagnostic evaluation programs are under development, one for mathematics at the elementary level, and another focusing on students' learning and communication processes across the curriculum in language arts, social studies and science at junior and senior high school levels. The diploma evaluations program evaluates students' achievement at grade 12 in English, social studies, mathematics and the sciences. Marks on this examination count for 50% of the final mark in these subject areas. The diploma examinations program is intended to help teachers and administrators assess the effectiveness of their instructional programs. The written response examinations, part A in English 30, and English 33 integrate reading and writing. English 33 examination also incorporates viewing. Both include personal responses to literature. The reading examinations, part B in both English 30 and 33 are literature based, providing full text and context for the multiple choice questions. English 33 also provides opportunities to demonstrate their editing skills and

their ability to select and use appropriate resources to solve practical problems. All science examinations include a written response section. The questions present novel problems and provide opportunities for students to demonstrate their communication skills and their abilities to engage in higher level thinking related to the important concepts of the course. For some of the computational questions in the math examinations students must work out their own answers which are recorded on a machine scannable answer sheet. This eliminates the guessing factor of multiple choice questions and results in very high discrimination factors between the top and bottom groups of students. Under development are guidelines to assist teachers in the assessment of students' problem solving and process skills in mathematics and the sciences. Also under development are materials that will assist teachers in evaluating students' oral language proficiency in a variety of situations. Materials submitted included a grade three achievement test, part A, writing and B, reading, grade 6 language art achievement test, part A, composition, and part B, reading, samples of student writing from the grade six English language arts achievement test, the 9th grade English language arts achievement test, part A composition and part B, reading, general information bulletin describing the achievement testing program for the school year 1989-90, description of the diagnostic communication processes integrating evaluation and instruction, and examples of the diploma examination in English, mathematics social studies, and several areas of science.

Comment: Alberta seems to have a very highly developed assessment program that mixes more traditional with authentic methods. There is clearly an emphasis on authentic assessment in these materials.
This is apparently avery ambitious and high quality program that should bear further fruit as it becomes fully developed.

ORGANIZATION: Albuquerque Public Schools

CONTACT PERSON: Dr. Sandra O'Neal, Acting Director
930-A Oak St. SE
Albuquerque, NM 87106

PHONE: 505-764-9711

ABSTRACT: Albuquerque Public Schools has developed alternative assessments in four key areas: 1)the Kindergarten Developmental Progress Report (KDPR); 2)the Word Recognition Placement Inventory (WRPI); 3)key competencies in mathematics assessment; and 4)high school subjects tests. Each of these was locally developed in response to a specific assessment need in the Albuquerque Public Schools. The KDPR provides kindergarten teachers with activities through which student progress can be measured for social, emotional development, physical development, and intellectual development. Observations can be transferred to a parent report or to a NCS form for computer processing. Student progress is rated as S = areas of strength, G = growth and progress, N = beginning/needs more time and experience. NA = not assessed or introduced at this time, A = area of special need/concern. The WRPI is used in grades 1 to 5 at the beginning of the school year for placing students at appropriate reading levels. The WRPI consists of graded lists of words for students to read individually to the teacher. The Key Competencies in Mathematics assessment is used in grades 1 to 8 to assess math competencies at each grade level. An activities based math assessment is under development for use with small and large groups of students. Key Competencies at grades 3, 4, and 5 are assessed with a combination of a written test and activity based assessments. At grades 6, 7 and 8, Key Competencies assessment consists of multiple choice tests. The High School Subjects Tests are a series of multiple choice tests designed to assess the high school curriculum in the areas of algebra, biology, U.S. History, 10th grade English and health.

ORGANIZATION: American Council on Education
GED Testing Service

CONTACT PERSON: Janet Baldwin, Senior Research Associate
1 Dupont Circle, West
Washington, D.C. 20036

PHONE: 202-939-9490

ABSTRACT: ACE's GED Testing Service has, since 1988, included an essay test in the writing skills portion of the GED. It is holistically scored. A technical paper, entitled "Adding an Essay to the GED Writing Skills Test: Reliability and Validity Issues" by Richard Swartz, Wayne M. Patience, and Douglas R. Whitney, was submitted. The paper reports on a study undertaken by the GED Testing Service to determine whether acceptable reading reliabilities could be attained through decentralized holistic scoring on a large scale, whether consistent standards could be applied by different individuals in different locations and whether an essay would measure aspects of writing performance that differed from that measured by a multiple choice test. Over 6,000 students participated in the study. Results indicated that acceptable levels of reading reliability are possible in decentralized scoring sessions. Readers awarded the same or contiguous scores for 93% of the essays. This rate matched the typical, centralized scoring sessions and yielded a reading reliability of 0.82. Additional research is planned to identify the most effective methods for maintaining scoring scale stability in essay scoring sessions conducted in a variety of settings.

ORGANIZATION: Anchorage School District

CONTACT PERSON: Ray Fenton, Ph.D., Research Associate
Assessment and Evaluation
4600 DeBarr Avenue
P.O. Box 196614
Anchorage, AK 99519-6614

PHONE: 907-269-2211

ABSTRACT: Anchorage School District reports using a variety of measures to assess student achievement. The effort to develop and improve measures of physical, intellectual, and affective items continues. In whole language assessment, they indicate that revision of ASDCRT tests and the development of portfolio assessment. Follow-up would be required to obtain materials.

ORGANIZATION: Baltimore Public Schools

CONTACT PERSON: Dr. Denise G. Border, Assistant Superintendent
Division of Planning, Testing, Research and Evaluation
Room 203
200 E. North Avenue
Baltimore, MD 21202

PHONE: 301-396-8948

ABSTRACT: This District reports developing a three to four year plan to produce a criterion referenced and performance based assessment program to replace the current norm-referenced testing program. No materials were submitted.

ORGANIZATION: Boston Public Schools

CONTACT PERSON: Maryellen Donahue, Director
Office of Research and Development
26 Court Street, 8th Floor
Boston, MA 02108

PHONE: 617-726-6200, ext. 5840

ABSTRACT: Performance assessment materials in reading, writing and mathematics are being developed for the BPS by Boston College. They are also pilot testing mathematics performance assessments developed by The Psychological Corporation. In a memorandum to zone superintendents, headmasters, principals and other administrative heads, the deputy superintendent describes Boston's assessment program. The memo states that, "The Center for the Study of Testing, Evaluation and Educational Policy at Boston College will design school assessment programs that will:
- help schools set clear goals on what they want students to learn
- monitor student progress toward these goals
- improve instruction by placing emphasis on higher order thinking skills instead of rote recall
- provide parents and community with accurate information about each school's progress toward its academic goals."

The assessment program includes the following features:
- The school site council will collaborate in setting goals and reviewing proposals to measure progress toward meeting them
- Portfolios, collections of student work, especially those that show the evolution of a final product will be used to measure student and school progress
- Progress will be assessed using a variety of indicators. In most cases, students will be required to produce a product or demonstrate the appropriate application of a skill or principle in solving a problem.

Sampling procedures will be used to limit the number of assessment activities individual children are subjected to and thus to increase significantly the range of skills and behaviors assessed. Boston College will pilot test these new assessment procedures and instruments in the areas of reading, writing and math in twelve schools; two elementary schools, and one middle school in each zone, and three high schools in 1990-91. This study will be expanded in the 1991-92 school year. The 1990-91 pilot studies' goals are to:
- develop assessment instruments and procedures in reading, writing and math in grades 2, 5, 8 10 and 12

- make the testing procedures and results more useful to teachers and administrators
- design appropriate and useful recording and reporting schemes
- help examine what information the general public needs
- explore various formats for reporting such information
- design a process that will enable individual schools to improve goal setting procedures and
- design training protocols.

Criteria for school selection are included in the memorandum and applications are solicited.

The contact person at Boston College is John E. Cawthorne - CSTEEP, Boston Assessment Program, McGuinn 531, Boston College, Chestnut Hill, MA 02167.

Comment: By the time these abstracts are published, the materials for this newly developed assessment program should be available. Follow up with Boston College would be advisable.

ORGANIZATION: CTB-McGraw Hill

CONTACT PERSON: Ellen Haley, Director
Strategic Product Development
2400 Garden Road
Monterey, CA 93940

PHONE: 408-649-7757

ABSTRACT: At the time of the survey, CTB was planning the development of effective, appropriate performance assessments to complement their multiple choice tests. They were investigating solutions to problems associated with performance assessment such as reliability, validity, scaling, cost, testing time, rater training, etc. CTB believes that rigorously developed performance assessments along with multiple choice tests can be used together to provide more information about student performance. Whole language assessment is addressed through the integration of reading, writing, and language assessment into broad based tasks that will measure multiple skills and multiple outcomes for grades 1 through 8. No materials were submitted.

ORGANIZATION: California Department of Education

CONTACT PERSON: Dr. Dale Carlson, Director
California Assessment Program
California Department of Education
721 Capitol Mall
Sacramento, CA 95814

PHONE: 916-322-2200

ABSTRACT: At the time of the survey, California was in the process of developing an integrated English language arts performance assessment and performance assessments in math, science and history. In the whole language area they were developing an integrated English language arts assessment that would assess reading performance through open-ended exercises and writing through essays in many types of writing. Materials submitted with this response include a 1987-88 annual report on writing achievement of California 8th graders, year 2, and a circular on writing assessment entitled "Thinking Through Writing". The circular describes California's writing assessment process including the types of writing assessed, prompts used, the scoring system and reporting systems used, the types of assistance provided, and future plans. It also contains sample essays by students who have taken the assessment. At the eighth grade, eight types of writing were assessed. They included autobiographical incident, evaluation, problem solution, report of information, first hand biography, fiction story, observational writing and speculation about causes and effects. At grade twelve, the 1990 assessment included autobiographical incident, evaluation, interpretation, reflective essay, controversial issue, report of information, observational writing, and speculation about causes or effects. In each case students were given 45 minutes to write an essay on the type of writing for which they were being tested. The scoring system is based on three dimensions: rhetorical effectiveness, a general feature such as coherence elaboration or style, and the use of conventions, i.e. usage mechanics and spelling. Each of these three scores uses a rating system of 1 through 6, teacher scorers are rigorously trained in the use of the scoring system to produce accurate and consistent scores.

Comment: California has one of the more advanced efforts in statewide application of authentic assessment activities. They are known for combining high technical standards with innovative approaches to large scale assessment. Dr. Carlson stands out as one of a handful of pioneers in innovative approaches to statewide assessment.

ORGANIZATION: Center for School Accountability

CONTACT PERSON: Dr. Roy H. Forbes, Director
Center for School Accountability
400 Oberlin Road, Ste 220
Raleigh, NC 27605

PHONE: 919-832-1584
919-334-5100

ABSTRACT: The Center for School Accountability (CSA) is encouraging districts to include optional indicators in the school improvement plans required by North Carolina's Senate Bill 2. Working with the task force on excellence on secondary education to include strong recommendations on changing the current state testing program to include "MAEM" techniques, whole language assessment is incorporated in the MAEM process. The only material submitted was a brochure on the Center for School Accountability describing the background, philosophy and primary activities of the center. Those activities are: 1) conducting research projects that advance the ability of schools to effectively and fairly assess student performance, 2) providing policy makers with workable methods for monitoring the quality and improvement of educational programs, 3) providing accountability related technical assistance to public schools systems and 4) developing instruments, procedures, and materials. Note: no specific assessment materials or research articles were included with this submission.

ORGANIZATION: Cherry Creek Schools

CONTACT PERSON: Gerald W. Bracey, Ph.D., Director
Research and Evaluation
Cherry Creek Schools
4700 South Yosemite
Englewood, Colorado 80111

PHONE: 303-773-1184, ext. 4233

ABSTRACT: The Cherry Creek District reports that they are developing a writing assessment, the assessment will incorporate elements of the writing process and will not be just a prompt. No materials were submitted.

ORGANIZATION: Cincinnati Public Schools

CONTACT PERSON: Deanna D. Spring, Supervisor
English/Communication Arts Mathematics
Cincinnati Public Schools
230 E. 9th St.
Cincinnati, OH 45202

PHONE: 513-369-4736

ABSTRACT: Cincinnati reports they are engaged in collection and assessment of writing portfolios for grades 1 through 12. Portfolio concepts will also become part of the mathematics assessment program as this is developed. In the whole language assessment area, they are using complete short stories, poems and selections from content texts in math, social studies, and science to assess reading processes, especially at grades 4-8. The emphasis is on process rather than content. No materials were submitted.

ORGANIZATION: Charlotte Mecklenberg Public Schools

CONTACT PERSON: Dr. Susan E. Henry, Pupil Assessment Specialist

P.O. Box 30035
Charlotte, NC 28230
704-379-7055

CONTENT/SUBJECT AREA: Reading, Writing

ABSTRACT: This district has budgeted a position for development of a holistic reading assessment instrument to measure student mastery in reading for grades 3, 4, 5 and 6; and for measuring mastery of competency goals in reading as outlined in the North Carolina Standard of Study. Writing assessments will also be developed.

An end-of-course test for English I will be implemented in 1991-92. It will require each student to write two essays. Four types of writing will be measured each year, expository, narrative, argumentative, and descriptive. Each year there will be a core prompt with three or four variable prompts. The scoring criteria will be main idea, supporting detail, organization, and coherence. These same criteria are used for grades 6 and 8. There will be a six-point focused holistic score for content and an analytic score for grammar. Students will analyze literary works using terms and concepts from the ninth and tenth grade curricula. Individual and summary scores will be returned on the core prompt but only summary scores for schools and school systems will be returned on the variable prompts. Testing will occur in March or April to allow for processing and reporting before the end of the school year.

ORGANIZATION: Colorado Springs Public Schools

CONTACT PERSON: Joe B. Hansen, Executive Director
Planning, Evaluation & Measurement
Colorado Springs School District #11
1115 N. El Paso St.
Colo. Spgs., CO 80903

ABSTRACT: Colorado Springs District 11 has developed a three point comprehensive assessment plan that incorporates objective referenced, norm-referenced and performance based measures. At the time of the survey, the district was working with the Northwest Evaluation Association (NWEA) to develop objective referenced measures in math, reading and language arts to meet state proficiency assessment requirement at grades 4, 8 and 12. This district was also planning to develop pilot projects in portfolio assessment in the spring 9f 1991. A copy of the three point plan was submitted.

ORGANIZATION: Connecticut State Department of Education

CONTACT PERSON: Douglas A. Rindone, Ed.D., Chief
Division of Research, Evaluation and Assessment
Bureau of Evaluation and Student Assessment
Connecticut State Department of Education
P.O. Box 2219
Hartford, CN 06145

ABSTRACT: Connecticut's common core of learning expresses the outcomes of education in the public schools identified by the Common Core of Learning Committee for the general purpose of improving the quality of the educational experience for all Connecticut students. "We believe there is a common set of skills, knowledge and attitudes essential to the total development of all Connecticut students." The Common Core establishes a vision of what Connecticut's high school graduates should know and be able to do. It represents a broad array of outcomes that should result from the entire K-12 school experience including academic skills, knowledge, personal and social skills, attitudes and attributes. It articulates Connecticut's expectations for its schools and its youth, thereby enhancing the quality and equality of educational opportunities throughout the state. One document titled, "Toward a New Generation of Student Outcome Measures: Connecticut's Common Core of Learning Assessment", is an AERA paper presented in San Francisco in 1989 with Joan Boykoff-Baron as the primary author. The paper describes, in a fairly comprehensive fashion, the assessment development project of the Common Core of Learning. If focuses on the experience gained within the project of developing performance assessment i science and mathematics. The paper sets forth guidelines for development and use of Common Core of Learning Assessments in grades 9 - 12. These guidelines include the rationale for common core of learning assessments, specifications for performance tasks, specifications for scoring rubrics and early development work.

The Connecticut Department of Education is a partner in the COMPACT which stands for Connecticut Multistate Performance Assessments Collaborative Team. This is a multi state collaborative effort funded by NSF in which Connecticut State Department of Education is collaborating with the Coalition of the Essential Schools and with the state departments of education in Michigan, Minnesota, New York, Texas, Vermont and Wisconsin for the purpose of developing performance tasks in math and science for high school students. These tasks are to be used by the CSDE to assess a statewide sample of students in 1991-92. COMPACT

produces a monthly newletter, "The COMPACT Dialogue", several copies of which are included in the submission. The submission also includes detailed examples of performance tasks developed according to specific guidelines.

Comment: This is among the most promising so-called authentic assessment projects in the country. It is widely reported and written about and some of the key individuals involved in the project such as Joan Boykoff-Baron, are well respected and known for their rigorous and thorough approach to the problems posed by authentic assessment. This is a project that should be closely monitored for new ideas and solutions as they are developed.

ORGANIZATION: Council Chief State School Officers (CCSSO)
State Education Assessment Center

CONTACT PERSON: Rolf Blank, Project Director, CCSSO
State Education Assessment Center
400 North Capitol Street
Washington, D.C. 20001

PHONE: 202-624-770

CONTENT/SUBJECT AREA: Science

ABSTRACT Materials submitted consist of a report on alternative methods for assessing science, dated February, 1989. The report summarizes results of a conference of state assessment directors and state science supervisors held in January, 1989. The conference was organized by the CCSSO Science, Math Indicators Project and sponsored by the National Science Foundation. It was authored by Rolf Blank and Ramsay Selden, both of the Council of Chief State School Officers.

The objectives of the conference were to: 1) increase the knowledge of state science supervisors and assessment directors of recent experience at international, national and state levels with alternative methods of student assessment in science and 2) to inform and assist states in planning alternative methods for state science assessment programs. A longer-term goal was to increase the number of states using alternative methods in science assessment. The report discusses recent developments in science assessment including research from ETS which tested hands-on science exercises and analyzed their potential application to large scale assessment programs as well as initiatives by NSF and the U.S. Department of Education.

The report indicates that 29 states have science assessment programs and many others are considering how to design science assessments to be more consistent with state objectives for curriculum and instruction. These authors indicate that 38 states have a curriculum framework or standards for science education that is used for selecting or recommending textbooks and guiding school curriculum and teacher inservice, with many states emphasizing, "science as a process" the authors suggest that state assessment programs consider hands-on science exercises as an approach to assessment. The report summarizes the conference proceedings including presentations from Senta Raizen, from the National Center for Improving Science Education, Rodney Doran, Professor of Science Education at the State University of New York at Buffalo,

reporting on the Second International Study, Walter McDonald, Science Assessment Coordinator for the National Assessment of Educational Progress (NAEP); Joan Baron, Director of the Connecticut Assessment of Educational Progress; Douglas Reynolds, State Science Supervisor for New York; Edward Roeber, Assessment Director for the State of Michigan; Zach Taylor, Science Consultant for California.

Key issues raised during the course of the conference concerning the development and use of alternative methods of science assessment included: 1) validity, 2)the role of assessment to drive vs. reflect the curriculum, 3) time requirements, 4) cost, and 5) use for trend data. Each of these points is discussed in the report.

In addition to the report summarizing the science assessment conference, the respondent also indicates that CCSSO is proposing to develop an interstate consortium on student assessment. He also indicates that CCSSO is leading a consensus process for assessment objectives for NAEP in reading, which promises to be the first state-by-state reading assessment. A contact person for reading objectives is Barbara Kapinus, Project Director, NAEP Reading Objectives, CCSSO State Education Assessment Center, 400 N. Capitol Street, Washington D.C. 20001, (202) 624-7700.

ORGANIZATION: Education Commission of the States

CONTACT PERSON: Beverly Anderson or Jane Armstrong
Director of Policy Studies
707 17th St., Ste. 27
Denver, CO 80202

PHONE: 303-299-3631

ABSTRACT: The Education Commission of the States (ECS) is working with six states to make changes in schools through local districts and states to change curriculum and instruction and assessment so that learning is more in-depth and focused on thinking and includes authentic measurement. ECS also helps to sponsor the annual Boulder Assessment Conference which has included many presentations on authentic assessment. No materials were submitted with this response.

Comment: ECS tends to function more as a facilitator of research and policy development rather than as a developer or researcher in its own right.

ORGANIZATION: School District 65, Evanston, Illinois

CONTACT PERSON: Judith Levinson, Director
Research and Evaluation
School District #65
1314 Ridge Avenue
Evanston, IL 60201

PHONE: 708-492-5950

ABSTRACT: This district reports that they are responding to state mandates for local assessment of Illinois goals and objectives. They are beginning to look beyond paper and pencil tests for students. Currently they have a writing assessment program in place using writing samples. Materials submitted include checklists in reading, speaking and listening.

ORGANIZATION: Fairfax County Public Schools

CONTACT PERSON: Lew Pike, Test Development Specialist
7423 Camp Alger Ave.
Falls Church VA 22042

PHONE: 703-698-0400

ABSTRACT: Fairfax County reports that they are using math CRTs for levels k- algebra that employ, in addition to regular multiple choice items, students' direct work and teacher observation. They also are using writing portfolios that are school based. They are engaged in development work with teachers to assess a new integrated language arts curriculum (Multiple Methods). They also mention school based resource and technical assistance.

Materials submitted include an examiner's manual for a mathematics test for primary level that includes items in a machine scorable booklet form. A higher level of a test reveals the inclusion of questions that require students to write story problems. The submission also includes a booklet of suggestions for assessment and evaluation in the integrated language arts curriculum. This is a handbook type document to be used by teachers. It includes instructions on how to set up a portfolio system, how to collect anecdotal records, conduct interviews with students, develop activity records such as individual conferences, language arts activities, reading, writing, strategies, lists, a personal writing record, and personal reading record, as well as some materials designed for student use. The manual contains reproducible masters for teachers to use in conducting the various forms of assessment addressed.

Comment: These materials reflect a high level of sophistication combined with simplicity of use. They are targeted directly at teachers and appear to be very utilitarian in design. This district should be included in any planned follow up on authentic assessment.

ORGANIZATION: Fillmore Unified School District

CONTACT PERSON: Chuck Weis, Assistant Superintendent
P.O. Box 697
Fillmore, CA 93015

PHONE: 805-524-0280, Ext. 26

ABSTRACT: This district is developing performance assessments in language arts for K-12 to assess progress toward expected student outcomes at grades 2, 5, 7 and 10. The assessments use structured observation, scoring rubrics and structured performances. At the time of the survey they were being field tested with students. In whole language, they have developed expected student outcomes in four language arts areas with a whole language orientation and accompanying assessments. These were in draft form. No materials were available for review.

ORGANIZATION: Georgia State Assessment

CONTACT PERSON: Dr. Linda Hansche, Director
Georgia Assessment Project
Georgia State University
Box 858
Atlanta, GA 30303

PHONE: 404-651-4230

ABSTRACT: Dr. Hansche is a member of an independent committee attempting to make a statement about assessment and reading, i.e., how not to do it as well as how to do it. They are addressing large scale assessment in the state of Georgia. No materials were submitted.

ORGANIZATION: Glenn Ellyn School District

CONTACT PERSON: Mrs. Bobbie Roeske, Test Coordinator
School District 89
789 Sheehan
Glenn Ellyn, IL 60137

PHONE: 708-469-8900

ABSTRACT: This district is using a district developed criterion-referenced testing system to test students in math, reading, science, social studies, and language arts. The CRT is based on a Rasch Model. This system makes use of the SAMS computer software developed by Michael Linacre, University of Chicago. SAMS is an item bank management system that provides electronic scoring and reporting.

ORGANIZATION: University of Hawaii

CONTACT PERSON: Dr. Morris K. Lai, Director
Evaluation Office
Curriculum and Development Group
College of Education
University of Hawaii
Honolulu, HA 96822

PHONE: 808-948-7900

ABSTRACT: The CRDG has, for several years, used and revised the Academic Learning Time Observation Coding Sheet, adapted from the Significant Bilingual Instruction Teacher's Study of Far West Regional Laboratory. They are developing a program implementation study model that scales the results of interviews, questionnaires, document review, and observations of attributes of programs, in mixed method design. No materials were submitted.

ORGANIZATION: Indianapolis Public Schools

CONTACT PERSON: Patricia Bolanos, Principal
Key School
1401 E. 10th St.
Indianapolis, IN 46201

PHONE: 317-226-4297

ABSTRACT: Indianapolis Public Schools has developed a magnet elementary school known as the "Key School" that is based on the concept of multiple intelligences as theorized by Howard Gardner (1983). The Key School has been the focus of articles in Education Week, The New York Times, and numerous other publications. This submission included a brief report entitled, "Connections, Animal Patterns, Changes in Time and Space and Interdisciplinary Theme Based Curriculum Report 1987 and 1988". According to the introduction "The actual mission of the Key School is to produce knowledge in the following areas: a. the role of the arts in general education; b. interdisciplinary curriculum development; c. collaborative professional development; and d. qualitative educational evaluation." All of these areas are integrated into the report. The evaluation strategy is based on individual student video portfolios containing student interviews and individual projects related to six themes suggested by Boyar as an appropriate structure for the nation's schools. Those themes are: the shared use of symbols, shared memberships groups and institutions; shared producing and consuming; shared relationships with nature; shared sense of time, and shared values and beliefs. A video archive of student portfolios is designed to fit a naturalist paradigm for qualitative educational evaluation. This archive will document the effects of this interdisciplinary theme based curriculum on each individual student. Student portfolios contain intitial interviews and the individual presentations by the student of his or her projects related to a particular theme. The report states that it is the intent of the school system to develop an assessment instrument with the video portfolios. Upon a student's departure from the Key School, a copy of the student's video portfolio, covering as many seven years of his or her development...(remainder of sentence cut off)

Materials submitted included copies of several articles about the Key School as well as an article on multiple intelligences.

ORGANIZATION: Jefferson County, Colorado, Public Schools

CONTACT PERSON: Diane C. Proctor, Coordinator
Program Evaluation and Testing
Jefferson County Public Schools
1829 Denver West Drive, Bldg. #27
Golden, CO 80401

PHONE: 303-231-2382

ABSTRACT: The respondent indicated that this district was still at the awareness and overview stage, doing introductory workshops and presentations in the spring of 1990. Some teachers have developed things they are using in their own classrooms but they are not of a high enough quality to disseminate. In the whole language assessment area, Jefferson County is "Tal;king about developing writing/reading portfolios that might include a variety of measures." They report that they have a writing sample and that they recently piloted a new format in which students were given two days and encouraged to engage in pre-writing activities. Reported highlights of the writing assessment program indicate coordination and integration of writing tasks into all subject areas. Writing is purposeful rather than practice; the importance of critical thinking skills is stressed as well as instruction in the writing process that is focused on teaching students how to think more effectively as they write. The writing test is administered at the middle of the school year in a single session of 45 to 50 minutes length in the elementary schools, and 1 class period in the secondary schools. Criteria for scoring are prescriptive based on mid grade level expectations and include the following components: ideas, style and vocabulary, organization and paragraph development, spelling, mechanical conventions, grammar and usage, sentence structure and handwriting/appearance. Scoring is holistic, performed independently by two specially trained graders. Examples of prescriptive criteria of the writing sample for grades 3, 5 and 7 are included in this submission. Also included are survey forms, one for teachers and another for secondary students, and for elementary students, used in evaluating the writing sample program.

ORGANIZATION: Kamehameha Schools

CONTACT PERSON: Katherine Tibbetts, Manager
Evaluation Department
CDEE, Kamahamaha Schools
Kapalaheina Heights
Honolulu, Hawaii, 96817

PHONE: 808-832-3067

ABSTRACT: Kamehameha Schools report they are using a portfolio system for assessment of language arts curriculum. The are also in the process of developing a system of performance based criterion-referenced assessment instruments for use by classroom teachers, program developers, program evaluators, and program administrators. No documentation was available at the time of our survey. A whole language approach is used in their language arts program.

ORGANIZATION: Los Angeles Unified School District

CONTACT PERSON: Linda Lownes, Assistant Director, or
Paula Moseley, Coordinator
Program Evaluation and Assessment Branch
450 N. Grand Ave. G-265
Los Angeles, CA 90012

PHONE: 213-625-6207

ABSTRACT: Los Angeles Unified School District currently uses a senior writing sample as part of their secondary competency program. This has been in place approximately twelve years. The Program, Evaluation and Assessment Branch is conducting background research on performance based assessment in an attempt to answer six questions about performance based assessment: 1)how does the literature define performance based assessment and authentic assessment; 2)what are the pros and cons; 3)what are the advantages and limitations; 4)what instruments have been commercially developed; 5)what instruments are currently used by other school districts; and 6)how can longitudinal data be reported? The PEAB contracted with Dr. Lamar Miller of NYU to gather the most current information about alternative student assessments. By interviewing experts from state departments of education, large city districts, researchers, superintendents, test publishers and measurement specialists. The PEAB is formulating plans to 1)begin developing a performance based algebra assessment; 2)continue to collect relevant research; 3)continue to talk with publishers; 4)investigate funds and resources available within the district.

The senior writing sample is one of four competency tests developed in response to a California State requirement for proficiencies certification for high school diplomas. Each student must complete two writing samples in a 50 minute period in response to descriptive prompts. Scoring is done holistically by two independent readers using a six-point scoring rubric. Schools receive pass-fail student lists each time the test is given. Numerous retest opportunities are provided for students who fail.

ORGANIZATION: LRDC, University of Pittsburgh

CONTACT PERSON: Suzanne Lane, Quasar Assessment Coordinator
Assistant Professor, Research Methods
717 LRDC, University of Pittsburgh
3939 O'Hara Street
Pittsburgh, PA 15260

PHONE: 412-648-8801

ABSTRACT: Dr. Ed Silver of the LRDC is directing a national demonstration project called "QUASAR" (Quantitative Understanding Amplifying Student Achievement and Reasoning) sponsored by a $10 million grant from Ford Foundation. It is intended to demonstrate that it is feasible and responsible to teach enhanced mathematics to middle school students in low income communities. Assessment instruments will be developed to measure students' growth in higher level thinking and reasoning skills in mathematics. Materials submitted were comprised of a project summary of the QUASAR project which includes the rationale for the project as well as an operational plan. The project began in May of 1989 and was expected to run through December of 1994. By September 1 of 1990, QUASAR is expected to have assembled a library of exemplary materials in assessment.

Comment: this project warrants further review and monitoring as it develops over time. Potentially, it could provide a new approach to assessment.

ORGANIZATION: Maine Department of Education

CONTACT PERSON: Paul R. Walker, Director, Division of Assessment
Station Norse Number 23
Augusta, ME 04333

PHONE: 207-289-5991

ABSTRACT: Maine's state assessment includes open ended and extended answers in reading, math, writing, social studies, science and humanities. Reading contains long passages which students are likely to enjoy reading. Whole language assessment emphasizes higher order skills and comprehension, i.e., there are no phonics on assessments at grade four and beyond. Materials submitted included aguide to the Maine Educational Assessment, 1989-90, published by the Department of Education and Cultural Services, Division of Educational Assessment. The guide describes program goals, schedule and test administration, participants, how the tests are developed and information on content and use. The Maine Educational Assessment (MEA) school reports include a diagnostic display where school scores within each major subject are reported. Data are displayed in a way that enables a local district to compare its standing against other districts within the state in an area or sub area of the objectives.

ORGANIZATION: Maryland State Department of Education

CONTACT PERSON: Robert E. Gabrys, Chief
Program Assessment Evaluation and
Instructional Support Branch
Maryland State Department of Education
200 W. Baltimore St.
Baltimore, MD 21201-2595

PHONE: 301-333-2369

ABSTRACT: At the time of the survey, the Maryland Department of Education had issued an RFP for the design, development and implementation of a state level criterion referenced testing program in mathematics, reading, writing and language usage at grades 3, 5, and 8. Proposals, according to the RFP, "must relate to the development of a battery of tests to include a mathematics and an integrated reading, writing, language usage component for each of the three grade levels." The deadline for proposals was April 13, 1990. This assessment program is a component of the Maryland State Performance Program, a 10 point program to improve education throughout the state of Maryland. The reading outcomes and test content specifications call for high quality thinking oriented measures. The RFP further indicated that the assessment battery should reflect and model process reading and writing in its format. It should use a reader-oriented interactive model for the reading component and a process writing for the writing component. Score reporting must address the issue of school improvement and individual student improvement. Detailed specifications are provided for each of the other components. By now the contract must have been awarded for the development of this program. Contact should be maintained with Robert Gabrys as the program is being developed.

ORGANIZATION: Michigan Department of Education

CONTACT PERSON: Edward Roeber, Supervisor
MEAP, Michigan Department of Education
P.O. Box 30008
Lansing, MI 48909

PHONE: 517-373-8393

ABSTRACT: The Michigan educational assessment program has a well-developed system of performance tests in art, music, science, social studies, physical fitness, mathematics and career development. They are also involved in whole language assessment. Materials submitted included: 1.) a technical paper on "Equating and the Use of Scale Scores" written by Peggy Dutcher and Susan Phillips; 2.) A position paper from the Michigan Reading Association entitled "Reading Redefined", which provides the background, rationale, a definition of reading, assumptions underlying the definition, and its implication for instruction and policy-making; 3.)an article from Educational Leadership, April, 1989, by Ed Roeber, titled "Michigan's Innovative Assessment of Reading", describing Michigan's reading test, its development, and proposed implementation; 4.)the essential skills reading blueprint of the Michigan Educational Assessment program. This is a very comprehensive document that not only provides the background, purpose, and rationale for the essential skills test, but also includes a description of the test, detailed information on item construction, test administration and reporting as well as examples of stories used at grades 4 and 7 and the informational selection process used at grade 10. This document is rich in detail and may represent the best single example of an actual performance assessment in use in the United States.

ORGANIZATION: Missouri Department of Education

CONTACT PERSON: Dr. William Boulter
Science Consultant
Missouri Department of Elementary & Secondary Education
P.O. Box 480
Jefferson City MO 65102

ABSTRACT: The Missouri Department of Elementary and Secondary Education reports that authentic science assessment instruments were in the beginning stages of development at the time of the survey, but no materials were available. Materials were submitted on Project Construct, a joint venture between the Center for Educational Assessment at the University of Missouri, Columbia, and the Department of Elementary and Secondary Education. Project Construct is described to be, "A process oriented curriculum and assessment framework for working with children aged three through seven." Construct is based on the view that children actively construct knowledge as a result of their interactions with the social and physical world, as opposed to developing through predetermined stages. Construct is a curriculum and assessment framework to support young childrens' unique ways of learning, while providing administrators, parents and teachers with the information needed to make appropriate decisions regarding the education of young children. The framework was developed by early childhood educators, administrators, subject area specialists and professors. Project Construct is based on the premise that children attain knowledge through interaction within four major domains: Sociomoral, cognitive, representational, and physical development. These four domains provide a framework for the development of curriculum and instruction materials. Goals and objectives have been developed within each of these four domains. The curriculum guide is intended to assist with rather than dictate curricular decisions to enable teachers to function as professional decision makers who use their knowledge and observations of young children to enrich teaching practice. The Construct assessment system will teachers document childrens' development. It will consist of an integrated set of evaluation tools referenced to the Project Construct goals for students. This assessment system will have two primary components: one to be used for formative evaluation and one to be used for summative evaluation. A variety of complementary instruments make up each components: those designed for formative evaluation are imbedded in every day classroom happenings and yield results that guide day-to-day instruction, while those designed for summative evaluation yield results that can be used to

summarize student performance and to determine which children may be eligible for special programs." Project Construct goals for students are intended to function as a bridge between the education parents provide at home as exemplified in the parents as teachers objectives and the core competencies and key skills identified for second grade, thus linking early childhood programs with the elementary curriculum. The goals and competencies are based on two common assumptions: that positive learning outcomes result from the alignment of curriculum instruction and assessment; and that all children can be successful in school.

Project Construct was expected to be available for all Missouri school districts and others working with young children to use on a voluntary basis beginning in the fall of 1990. Information about Project Construct can obtained from: Center for Educational Assessment, University of Missouri, Columbia, 403 S. 6th St., Columbia Missouri, 65211, (314) 882-4694; or Department of Elementary and Secondary Education, Division of Instruction, Early Children Education, P.O. Box 480, Jefferson City Missouri, 65102, (314)-751-(last 4 numbers missing from tape) subheading Missouri Writing Assessment, 1989. Included in this submission was a brief report on the results of the 1989 Missouri writing assessment which is also a joint venture of the Department of Elementary and Secondary Education and the Center for Educational Assessment at the University of Missouri, Columbia. The writing assessment is administered to representative samples of students in grades 6, 8, and 10 annually. This assessment uses a modified holistic scoring method based on a six-point scale. Each essay is scored by at least two readers. A third reader is used to resolve discrepencies. Data is reported at building level to provide comparisons to state results to grade level. The six-level scoring rubric includes five components at each level. They are: focus, organization, development, mechanics, sentence structure, and diction.

ORGANIZATION: Mount Diablo Unified School District

CONTACT PERSON: Robert P. Grobe, Administrator
Research and Evaluation
1936 Carlotta Drive
Concord, CA 94519-1397

PHONE: 415-682-8000, ext. 4112

ABSTRACT: The Mount Diablo Public Schools has developed an open-ended assessment of reading in grades three through eight and an open-ended assessment of mathematics in grades three through eight. Students are given a situation or problem, then asked to describe the procedure they would use to solve the problem. The material submitted was a report on curriculum based assessment for math, a summary of 1990 field test results. The report provides a)an introduction and rationale for curriculum based assessment. It is based on the precept that "in order to accurately assess students' achievement on some of the 'higher order' grade level expectations it is necessary to have students generate rather than select answers." The report states that, "testing of problem solving or the ability to apply concepts and skills in new and unexpected situations should include items that require students to formulate mathematical problems, select alternate mathematical strategies for solving problems, make generalizations and verify and interpret solutions." The report presents scoring information and results of the May 1990 curriculum based assessment in math field test.

Scoring was accomplished with a five point scale ranging from 0 to 4 representing competency levels wherein 0 = no attempt, 1 = incorrect response, 2 = inadequate response, 3 = satisfactory response, and 4 = exemplary response. The report provides guidance for relating assessment to classroom to classroom instruction. Four types of assessment are identified as critical to classroom assessment of student mathematical learning. They are: performance tasks, investigations over time, teacher observations, and portfolios. For each of these a list of instructional tips is provided in a tabular format. The remainder of the report presents results from the assessment as well as examples of student work.

Comment: This appears to be a unique approach, at least for an individual school district to take in assessing math performance. It has similarities to some of the techniques used in the Connecticut Consortium and those being developed and piloted in the California Assessment Program.

ORGANIZATION: Multnomah ESD

CONTACT PERSON: Peter Wolmut, Director
School Support Services
Multnomah ESD
11611 Northeast Ainsworth Circle
Portland, OR 97230

PHONE: 503-257-1770

ABSTRACT: Multnomah ESD, an intermediate education agency serving districts in Multnomah County Orgeon, conducts an analytic writing assessment for client districts. They also report having two pilot projects on portfolio assessments underway. The ESD conducts an analytic writing assessment using a six trait analytic model and provides scoring and reporting services. The six trait model is based on the traits of: 1) ideas/content; 2) organization; 3) voice (style); 4) word choice; 5) syntax; 6) writing conventions. Materials submitted include a description of the analytic trait model and descriptions and layouts of four different reports, an alpha class list, a student profile, a building summary, and a district summary.

Also included is a one page description of the pilot portfolio projects. The first is in the Gresham School District. This project is conducted in conjunction with a mathematics program that emphasizes language and communication skills as part of the math curriculum. Performance assessment and portfolio assessment techniques are used to assess goals associated with mathematics communication and problem solving skills. The second pilot program in the Reynolds School District is a portfolio assessment program for at-risk students. This project explores assessment models to allow the district to generate reports without placing requirements on what materials students and teachers may place into their portfolios.

The contact person for analytic writing is Mr. Barry Reinstein, 503-257-1773. The contact person for portfolio assessment is Dr. Leon Paulson, 503-257-1774.

ORGANIZATION: New Brunswick Department of Education

CONTACT PERSON: Daina Watson, Coordinator of Testing
Box 6000
Fredicton, New Brunswick E3B5H2, CANADA

PHONE: 506-453-2744

ABSTRACT: The New Brunswick Department of Education is involved in scaled measurement of interactive production of language as well as performance assessment in elementary and junior high school science. Whole language assessment is addressed through portfolio assessment and the writing assessment component of the Provincial Examination Program. Materials submitted outlined the series of recommendations for a grade eight pilot portfolio assessment project.

ORGANIZATION: Northwest Regional Eduactional Laboratory (NWREL)

CONTACT PERSON: Dr. Rick Stiggins, Director
Center for Classroom Assessment
101 S.W. Main St.
Portland, OR 97204

PHONE 503-275-9559

ABSTRACT: Dr. Stiggins is perhaps best known for his research in tne area of classroom assessment and for the development and provision of training programs for teachers and the development of high quality performance assessments for classroom use. In terms of whole language assessment, Dr. Stiggins's work provides teachers with assessment training directly focused on the evaluation of writing and reading proficiencies including strategies for the integration of assessment in the day to day classroom instruction. Materials submitted with this response incluce a brochure for the book: Creating Writers: Linking Assessment and Writing Instruction, by Vicky Spandell and Richard J. Stiggins, published by Addison Wesley/Longman. The book is a comprehensive treatise on developing writing assessment and integrating it with the instructional process. Also included is an article by Rick Stiggins describing and NCME instructional module on design and development of performance assessments. This article is from the ITEMS Series, published by NCME dated fall, 1987.

ORGANIZATION: Pennsylvania Department of Education

CONTACT PERSON: Leslie Peters, Assessment Specialist
Department of Education
333 Market St.
Harrisburg, PA 17126-0333

PHONE: 717-787-4234

ABSTRACT: At the time of the survey, Pennsylvania was conducting a pilot test on writing assessment as part of their statewide assessment program in the whole language assessment area. Their statewide reading assessment has been redesigned to eliminate testing for basic skills objectives and to instead measure students' ability to understand meaningful connected text. Materials submitted include two brochures; one describing reading tests for 1990 used by the Pennsylvania assessment system; the second describing the writing sample field test. For the writing sample field test, 10,000 grade six-nine students were asked in the spring of 1990 to produce a writing sample in response to specified topics or prompts. Samples were scored by groups of teachers, supervisors and administrators who were trained by the Pennsylvania Department of Education staff. These tests are entitled "TELLS Reading and Math Tests". TELLS stands for Testing for Essential Learning and Literacy Skills. The reading assessment is based on four elements: 1) prior knowledge about the reading passage's topic, 2) ability to construct meaning to comprehend the package, 3) ability to use reading strategies and 4) reading habits and attitudes. In order to assess a broad range of comprehension from literal to higher order thinking, test questions will address three levels of processing: 1) explicit, 2) implicit, and 3) extended. Explicit requires students to identify locate or confirm information directly stated in the passage, implicit requires the use of textual information and prior knowledge to construct meaning and draw conclusions, extended requires the student to respond to and think beyond the test.

Comment: Pennsylvania seems to be one of those states that has a rather promising approach to authentic assessment on a large scale. Further information on this program as it is developed and implemented should be collected.

ORGANIZATION: Pittsburgh Public Schools

CONTACT PERSON: Dr. Paul G. LeMahieu, Director
Research, Testing and Evaluation
Pittsburgh Public Schools
341 S. Bellefield Ave.
Pittsburgh, PA 15213

PHONE: 412-622-3942

ABSTRACT: Pittsburgh Public Schools reports on four developmental programs in authentic assessment currently underway. These programs are: 1)Monitoring Achievement in Pittsburgh (MAP); 2)Syllabus Examination Project (SEP); 3)Arts PROPEL; and 4)Science Portfolio Project. A summary of each is provided, as well as detailed descriptive materials for each. Monitoring Achievement in Pittsburgh attempts to consciously link curriculum instruction and assessment to routinely provide data describing students' progress for teaching and learning decisions. Numerous performance based assessments have been incorporated within MAP. For example, in mathematics, a variety of performance tasks are assigned in conjunction with each MAP administration. MAP includes both a composition and a critical thinking component, each of which requires students to write extended essays. The composition exercises are nested within a process approach to writing instruction. "In the critical thinking component, students respond to a prompt that requires them to evaluate information, establish a position, and present the strongest possible exposition of that position, while selecting and appropriately using relevant data and evidence."

The Syllabus Examination Project is Pittsburgh's initial effort at developing course specific examinations for all secondary academic courseware. At least 60% of the SEP examinations are performance based. These have evolved from extended essays to include practical task problem solving in mathematics, the demonstration of important concepts or other performances in science as well as portfolio assessments in English. Eventually this array of assessment tasks may be extended to inlude oral presentations and demonstrations, debate performances, other performances and extensive projects. SEP course syllabi are included in this submission.

Arts PROPEL, a collaborative effort of Harvard Project Zero, the Educational Testing Service, and the Pittsburgh Public Schools, explores a new alternative formulation for the organization and delivery of curriculum instruction and assessment. It emphasizes perception, reflection, and production in art and

defines each as appropriate to various content areas. "PROPEL also seeks to build new forms of assessment and make assessment a more integral aspect of teaching and learning." Extensive descriptive material is included in this submission.

Science Portfolio Project is a collaborative of the University of Pittsburgh, ETS, and the Pittsburgh Public Schools. This nascent project will explore the establishment of a "portfolio culture" and portfolio assessment of students' learning in the area of science.

Comment: Perhaps more than any other major school district in the United States, Pittsburgh has developed collaborative working relationships with major universities, research institutions and foundations to explore and develop innovative approaches to curriculum instruction and assessment. The materials provided here are of excellent quality, demonstrating the richness that has emerged from these collaborative efforts.

ORGANIZATION: The Psychological Corporation
Harcourt, Brace, Jovanovich

CONTACT PERSON: Dr. Eugene Jongsma, Senior Project Director
555 Academic Court
San Antonio, TX 78204

PHONE: 512-554-8108

ABSTRACT: The Psychological Corporation has developed an integrated assessment system for integrated language arts. This is a portfolio assessment approach to language arts based on seven components: 1) introductory video tape and viewer's guide which introduces teachers to portfolio assessment and provides instruction on how to start and maintain language arts portfolios; 2) a teacher's manual in a three-ring binder format with an orientation to portfolio assessment, guidance for record keeping and numerous forms and logs as well as suggestions for aggregating and using collected data; 3) individual student portfolio folders; 4) a portfolio storage box; 5) reading and writing activities including a variety of types of reading material both fiction and non-fiction intended to elicit a variety of higher order thinking skills; 6) scoring guides for each reading and writing activity; and 7) a scoring training package designed for principals, curriculum leaders and teachers to score the reading and writing activities. This contains a video tape and printed support materials.

This submission consisted of an advertising brochure, a set of sample materials was not included. However, such a sample kit could probably be obtained from a regional representative of the Psychological Corporation.

ORGANIZATION: Ministry of Education, Government of Quebec

CONTACT PERSON: Beverly Steele, Director of Development of Evaluation
600 Rhue Fullum 9E Etage
Montreal, Quebec, H2K4L1

PHONE: 514-873-2161

ABSTRACT: The Education Ministry of Quebec is developing a high school leaving examination characterized by a portfolio approach. Students have an opportunity to pre-write and collaborate. Quebec is very proud of its whole language English language arts curriculum. Whole language assessment is approached as a prototype alternate assessment in both written and oral communication. Extensive materials were included with this submission. The materials include six booklets on the secondary leaving examinations. The first is labeled "Administration and Invidulization Guide". It contains general information on administration, scheduling, and procedures for conducting the examination. Second is the Marking Guide for English Language Arts 516. This guide describes criteria and indicators to be used when correcting the Comprehensive Secondary School Leaving Examination in English Language Arts 516. Third is Student Booklet, Part 1. This booklet contains information about and instructions for the Secondary School Leaving Examination which has three parts. Part one contains in-class activities using the student booklet and the literature booklet. Part two consists of a major assignment to be written individually by each student. It draws on a portfolio of work compiled during part one and the literature booklet. Part three consists of a major assignment to be written individually. This also draws from the portfolio of work compiled during part one. Third is the Student Booklet, Part 2 containing general instructions for part 2 of the exam. Fourth comes Student Booklet Part 3, containing the general instructions for part 3. Fifth is the Literature Booklet containing 18 literary passages of poetry, short story, and essays for use by students in the assessment. Also submitted were two evaluation reports, a program evaluation report on English language arts secondary 2, dated Feb. 1990. This contains the results of survey and interview studies of parents, teachers, administrators, and consultants reactions to the ELA program. Positive results were found in a number of areas: 1) teachers believed that they had grown in their professional and personal development as a result of the program, 2) students felt that they benefited from collaborative encounters with their peers. 3) Consultants and teachers recognized the need for time to plan in an informed manner, while teachers and consultants appeared to have a very clear cut idea of what their needs are and how best to accommodate them. School principals appeared to have a realistic understanding of the program's requirements, while at the same time demonstrating an awareness of the need to provide the staff with professional improvement opportunities. Parents endorsed the language principles of the program, and made it clear that they feel that they are part of the school

community. Sixty six percent reported that their child had a positive attitude toward English language arts. Results of the student examination indicated a need to perfect the measurement tools used in the evaluation. Although all students in the sample produced narratives and seemed to know the essential attributes of the story form, too many students were observed to stop short of polishing and refining their writing. The results of the sample showed that students had difficulty practicing strategies for revising and proofreading. Recommendations were made for further research and for the refinement of instruments and materials as well as for improvements in inservice.

Comment: This appears to be a well-developed program, with an active and continuing evaluation component. It should rank among the better efforts of its type in North America.

ORGANIZATION: The Riverside Publishing Company

CONTACT PERSON: Fred L. Finch, Vice President/Editor in Chief
The Riverside Publishing Company
8420 W. Bryn Mawr Ave.
Chicago, IL 60631

PHONE: 312-693-0040, ext. 500

ABSTRACT: Riverside reports developing under contract to the Arizona State Department of Education, performance tests to assess reading, writing, and mathematics at grades 3, 8, and 12. Materials submitted consisted of an invitation to participate in a research study for the development of an integrated literature and language arts portfolio program. Riverside plans to include a new kind of assessment instrument that surveys a student's interests and experiences in reading, writing, speaking and listening. "It provides a way to assess a student's overall reading ability, including the ability to use reading strategies. It provides an assessment of a student's proficiency in the other language arts in the context of a reading for listening experience." The proposed integrated literature and language arts portfolio program appears to be a three-part assessment that includes a teacher checklist, the student interest and experience survey, and performance exercises. Two forms of a test each, for grades 2 - 8 will allow for pre and post testing. The test will need to be hand scored at the local level or by a scoring service that will be made available.

ORGANIZATION: Independent School District #196
Rosemount, MN

CONTACT PERSON: Jim Angermeyr, Coordinator of Program Evaluation
Independent District #196
14445 Diamond Path
Rosemount, MN 55068

PHONE: 612-423-7732

ABSTRACT: This District reports working to develop a science process test with emphasis on teamwork, cooperative problem solving, higher order thinking skills and ability to integrate the science process. They were planning to administer the test in the fall of 1990. The assessment was to include the following four components: a paper and pencil test of process science concepts containing one item for each of the required units in grades 3, 4, and 5, and two or three items on data use and interpretation. Second was a demonstration/hands-on test composed of twelve stations or circuits with a station for each of the required twelve units of the third, fourth, and fifth grade curriculum. A third component is to be designed after components on and two have been completed. It will test process skills and problem solving capability on one item that is totally unfamiliar to students. A fourth component will be comprised of actual observations of students as they work on the stations or circuits. The observer will concentrate on evidence of students' ability to perform the requested task, indications of confusion or lack of knowledge, comfort in working with apparatus, etc.

Comment: This sounds like an interesting excursion into a very difficult area of authentic assessment. A follow-up should be conducted in the spring of 91 to check on the progress of this project.

ORGANIZATION: Sacramento City Unified School District

CONTACT PERSON: Morris Mueller, Research Specialist
Sacramento City Unified School District
1619 N Street
Sacramento, CA 95814

PHONE: 916-553-4050

ABSTRACT: This district reports that it is developing a portfolio for grades 2 through 6 in language arts. Materials submitted include a set of recommendations from the Language Arts Assessment Committee specifying that teachers should establish a language arts folder/portfolio for every student and send on to the next year's teacher in grades 2 through 6 one set of student writing on one topic during the last quarter and one progress survey. Included in the submission are guidelines for student writing for grades K through high school, a scoring guide for holistic scoring of student writing samples and a writing guide for teachers indicating characteristics of student writers that may be seen at various grade levels. Also included are writing objectives for grades K through high schools and a document containing desired exit skills for writing at elementary, intermediate, and high school. The latter document is a product of the Clovis Unified School District in Clovis, CA.

ORGANIZATION: San Diego City Schools

CONTACT PERSON: Grant Behnke, Ph.D., Director
Testing and Consolidated Applications Dept.
Education Center Annex 3
4100 Normal St.
San Diego, CA 92103

PHONE: 619-293-8509

ABSTRACT: San Diego is in the process of developing a new testing system with a variety of components including performance based and open response questions. They're also in the early stages of working with teacher groups and curriculum experts in attempting to come up with valid whole language assessment activities.

Materials submitted included two documents relevant to authentic assessment. One, titled "The New Testing System Component Progress Reports", dated January 25, 1990, contains several brief reports that update the progress of various components of a proposed new testing system. The components include direct writing, and, of course, portfolio, computer managed assessment and computer adaptive testing. The progress report document delineates twelve key points of the proposed system. Collectively, these twelve points provide the basis for a comprehensive research development program to meet all current and anticipated assessment needs in San Diego. The first point, which is a key to the entire plan, calls for the elimination of district-wide testing as it has been conducted by consolidating it with the state-wide CAP program. The comprehensive plan for assessment in San Diego City Schools includes norm referenced, criterion referenced and performance based assessments; each appropriately targeted on different assessment needs.

Comment: Although the San Diego program is still in the planning stage, it warrants close monitoring since it is well conceived, coordinated with the California Assessment Program, is being developed in collaboration with the Psychological Corporation and is supported by a strong staff with a strong reputation for evaluation and assessment quality.

ORGANIZATION: St. Louis Public Schools

CONTACT PERSON: Jerry M. Powers, Director
Division of Evaluation and Research
St. Louis Public Schools
911 Locust
St. Louis, MO 63101

PHONE: 314-231-3720, ext. 346

ABSTRACT: St. Louis Public Schools conducts and annual writing assessment at four grade levels. They are alsc beginning to examine speaking assessment. This submission was a report on the writing assessment results for the spring of 1990 dated September 1990. The assessment targeted all students in grades 3, 5, 8 and a sample of sophomore English students. Students at grade 4 were also sampled in seven elementary schools and at grades 6 and 7 in four middle schools. The writing topic was designed to elicit narrative and descriptive writing. Students were asked to identify and write about a job that would be of interest to them. Papers were scored holistically by groups of teachrs appropriate to each level. Each paper received a numeric score based on an overall single impression of the composition. A six-point rubric with 1 representing the lowest possible score and 6 the highest was used for scoring. Two raters scored each paper. If scores for a given paper differed by more than 1 point a third rater was used to resolve the discrepancy. The report contains a detailed explanation of the scoring rubric, examples of papers from each level, and a summary of results. It is a well written, well documented report.

ORGANIZATION: South Carolina Department of Education

CONTACT PERSON: Dr. Marsha DeLain, Education Program Specialist
1429 Senate Street
Rutledge Building, Room 607
Columbia, SC 29201

PHONE: 803-734-8266

ABSTRACT: The South Carolina Department of Education was, at the time of the survey, in the process of reviewing their statewide criterion referenced instruction and assessment program. As part of this process, they were considering more authentic assessment in several areas. Future assessments in mathematics may allow students to use mathematic tools including calculators on certain portions of the test. If financially feasible, in math and reading, some of the items may require a constructed response, as is currently done with writing assessments. In early childhood, consideration was being given to limiting testing of young children (k-3). Whole language assessment may be assessed through a constructed response format for some test items. This may be accomplished by using a "writing in response to reading" approach to determine whether a student comprehends a reading passage. This approach would also serve as a bridge between the program's writing and reading components, hence promoting a total language approach. Material submitted was a report on a statewide survey of public school administrators and all K-12 teachers in the state. The survey contained sections on early childhood education, reading, mathematics and writing. Each of these sections consisted of proposed changes followed by a brief rationale regarding why each change was viewed as an important element to include in the basic skills assessment II (BSAP II). Respondents were asked to indicate on a 4- point scale whether they fully support, partially support, do not support, or had no opinion about each recommendation. Open ended comments were also collected. Results indicated very strong support for the proposed changes.

ORGANIZATION: Southwest Educational Development Laboratory

CONTACT PERSON: Dr. Ida Jean Holman,
Senior Training and Technical Assistance Associate
SEDL 211 E. 7th St.
Austin, TX 78701

PHONE: 512-476-6861

ABSTRACT: SEDL has promoted the concept of authentic assessment through audio tele-conferences, regional meetings, and a training of trainers workshop. A transcript of a November, 1988 teleconference with Dr. Barbara Presseisen of Research for Better Schools was included with the materials submitted. A teleconference in November, 1989 with Rick Stiggins of Northwest Regional Educational Laboratory (NWREL) focused on definition of teacher competencies for classroom assessment. Dr. Stiggins also conducted two workshops in Austin concerning "What Teachers Need to Know to Assess Student Performance", a component of which concerned the use of performance assessment techniques. On the assessment of whole language issue, this respondent indicated that SEDL has gone to the sites of schools in Arkansas and New Mexico where the whole language approach is being used. They have heard presentations about whole language from schools in Houston and Oklahoma and have looked at the issue of holistic scoring and its implications for measuring student achievement in a synthesis document on "Facilitating Student Achievement in Reading, Writing and Thinking With Partnerships and Technology". A chapter of that publication is included in the materials submitted. A cursory review of that chapter entitled "Writing as a Facilitator of Student Achievement" reveals that it includes a definition of writing and its purposes, identifies elements that influence effective writing, explores the relationship of writing to student achievement, summarizes ways of measuring students' writing skills and ways of measuring writing program effectiveness.

Comment: Although SEDL is not an institution that focuses on assessment, it seems to be at the hub of a lot of local activity involving performance assessment, holistic assessment, and other techniques. Continued contact with Dr. Holman would be advisable.

ORGANIZATION: Stanford University

CONTACT PERSON: Professor Robert Calfee
School of Education
Stanford University
Stanford, CA 94305-3096

PHONE: 415-723-8698

ABSTRACT: Professor Calfee reported that the work by he and his colleagues focuses on procedures that allow informed teacher judgment of student achievement to serve as a complement to standardized assessment techniques. Materials submitted include a paper titled "Oral Language Assessment Through Formal Discussion" by Robert Calfee and Lynn Sutter, both of Stanford University. The article deals with issues related to the development of oral language skills, various techniques for assessing oral language development, modifications of those techniques, and a focus on a discussion based approach to oral language assessment. A second article by Calfee, titled "The School as a Context for Assessment of Literacy" was published in The Reading Teacher, April, 1987. A third article is a chapter for the Handbook of Research on Reading, Volume II, entitled "Classroom Assessment of Reading", Robert Calfee, Stanford University, and Alfreda Hiebert, University of Colorado, Boulder. A review of section headings indicates that this chapter deals with literacy and assessment from a conceptual perspective, influences of externally mandated assessments, including effects on teachers' roles in curriculum and instruction, influence of internal assessment which describes an internal assessment model and techniques for teacher directed assessment, as well as effects on teachers' roles in curriculum and instruction. This article has an extensive list of references addressing many aspect of research on instruction and assessment. A fourth article, written for teacher assessment March 25, 1988, also by Hiebert and Calfee, is entitled "What Research has to say to the Classroom Teacher: Assessment for Instructional Decision Making." These articles and publications would seem worthy of further review for possible inclusion in a comprehensive literature review.

ORGANIZATION: Toronto Board of Education

CONTACT PERSON: Dr. Sylvia Larter or
D.C. Hambleton, Educational Research Coordinator
155 College St.
Toronto, ON M5T1P6

PHONE: 416-393-9957

ABSTRACT: Toronto has a highly developed a sophisticated system of performance assessments integrated with their curriculum in language arts and mathematics. This system, "Benchmarks", is a series of libraries containing both video and print materials. Each library contains a list of ministry goals and objectives upon which performance tasks are based. The actual tasks as performed by students and the criteria established for use in evaluation, graphs showing the range of student performance where applicable, samples of student work and descriptions of particular strengths and weaknesses. A brochure contains examples of language arts and mathematics tasks with criteria and instructions for holistic scoring of each. The Benchmarks are used to establish standards of achievement for the Toronto Board of Education, to provide system-wide reference points to assist teachers in student evaluation, to provide more comprehensive information for reporting to parents, to clarify the kinds of performance that lead to academic success, and to support teachers in their implementation of ministry and board objectives. The enclosed literature indicates that grades 3 and 6 language and mathematics Benchmarks will be available in November 1989, grade 8 Benchmarks will be available in fall 1990, and grade 10 Benchmarks are planned for the near future.

Comment: This appears to be a highly polished, commercial quality support system for teacher in performance assessment. An attempt should be made to acquire Benchmarks packages.

ORGANIZATION: University of Utah

CONTACT PERSON: Professor Gabriel M. Della-Piana
Professor of Educational Psychology
327 Milton Benyon Hall
University of Utah
Salt Lake City, UT 84112

PHONE: 801-581-7148

ABSTRACT: Professor Della-Piana has summarized his own work in the area of authentic assessment in a three point letter. First is teacher training in measurement. He has included a draft report on teacher competence in measurement and testing. He indicates that he has been following recent developments in an attempt to specify what measurement teachers need to know. Second, he advocates the writing conference as an approach to informal assessment. He has included two papers, one titled "Teacher Advisement Skill, Student Writing Skill, and Assessment Policy", the other, "Making Diagnostically Important Behaviors Observable in Writing Instruction", is a more recent and different elaboration of part of the first paper, and is an unpublished manuscript. Dr. Della-Piana mentions three other lines of work in which he is currently engaged, but for which no documents are available. First is a teacher dialogue with teachers where the assessment focus is a teacher problem solving with peers on a promise and problems of their own specific plans and implementations of writing instruction. Second is assessment in the context of whole language instruction in an outcome based education program. Third is assessment in the context of whole language student centered writing instruction.

Comment: Dr. Della-Piana is obviously involved in a wide range of authentic assessment related activities and should be included as an important source of information on new developments in authentic assessment. A detailed review of his papers should be part of any comprehensive literature review on this subject.

ORGANIZATION: Ministry of Education, Educational Programs Division
Victoria, British Columbia

CONTACT PERSON: Ann Chatwin and Nancy Procter, Research Officers
Student Assessment Branch
Ministry of Education
Educational Programs Division
Parliament Buildings
Victoria, B.C., V8V2M4

PHONE: 604-356-2418

ABSTRACT: The Victoria Ministry reports that they are developing oral and listening assessment in French language, interviewing techniques for assessment in math, performance assessment techniques to evaluate students in new primary, intermediate, and graduate program curriculums. In the area of whole language, they are developing teachers' resource packages for evaluating oral communication and written expression, developing resource material for evaluating thinking skills and they are refining holistic scoring procedures in history 12, communications 12, English 12, and English Literature 12, of Provincial Exams. The submission includes holistic scoring procedures for Provincial of scholarship examination essays for English Literature 12, holistic procedures for communications 12, and for English 12. Also included is a teacher's resource package on enhancing and evaluating oral communication in the secondary grades. This handbook for teachers includes instructional strategies in listening, and speaking as well as discussion groups and debating. It also includes evaluation strategies with sample planning and evaluation guides and student evaluation forms.

Comment: These are quite sophisticated and yet very practical performance assessment tools in the language arts. They appear to be consistent wtih the outstanding quality of the other Canadian efforts reviewed thus far.

ORGANIZATION: Virginia Department of Education

CONTACT PERSON: Mr. Donald Sale, Supervisor of Testing
P.O. Box 60
Richmond, VA 23216

PHONE: 804-225-2914

ABSTRACT: Virginia has in place a Standards of Learning Program comprised of learner objectives in all subjects in all grades from K-12. The whole language approach is used in the primary grades. Submitted materials included Objectives in the Language Arts with specific instructions including teaching suggestions, assessment strategies, student performance examples, scoring instructions, a sample checklist for performance rating, and score examples. This appears to be a well developed program used throughout the state of Virginia.

ORGANIZATION: Washington D.C., Public Schools

CONTACT PERSON: Dr. Cynthia Almeida, Assistant Director
Student Assessment
D.C. Public Schools, Bryan Building, Rm 206
13th and Independence Ave. SE
Washington, D.C., 20003

PHONE: 202-727-2092

ABSTRACT: The D.C. Public Schools Student Assessment program includes writing assessment in grades 4, 7, and 10 and production items in the end-of-course examination program at secondary level. Whole language is addressed through writing assessment and essays at the secondary level in the end-of-course testing program.

ORGANIZATION: None

CONTACT PERSON: Esther E. Diamond, Ph.D.,
Educational and Psychological Consultant
721 Brown Ave.
Evanston, IL 60202

PHONE: 708-328-7043

ABSTRACT: This consultant was involved in the development of Code of Fair Testing in Education. This is an activity of the joint committee on testing practices (NCME-AERA), a working group on testing to increase public awareness and understanding. Materials submitted consisted of a journal article titled "Fair Testing Practices: Issues and Resolutions, The Work of the Joint Committee on Testing Practices," Esther E. Diamond, Educational and Psychological Consultant, published in Social Applications and Issues in Psychology, R.C. King and J.K. Collins, Eds. Elsevier Science Publishers, BV North Holland, 1989.

AUTHENTIC ASSESSMENT SURVEY ABSTRACT

ORGANIZATION: Vermont Department of Education

CONTACT PERSON: W. Ross Brewer, Director of Planning and Policy Development

Vermont Department of Education
120 State Street
Montpelier, VT 05602
802-828-3135

CONTENT/SUBJECT AREA: Writing, Mathematics

ABSTRACT: Vermont assesses the writing and math achievement of fourth and eighth grade students using three methods: a uniform test, a portfolio and a "best piece." Thus combining traditional and alternative assessment techniques.

All grade 4 and 8 students take a test that uses equivalent tasks administered under the same conditions. In the 1990-91 pilot year, a uniform test will be administered in writing only. In the following year students will take a uniform test in both writing and mathematics. This test will be a multiple choice format and will include items from the NAEP, which will allow for comparisons of Vermont's students with those of other states.

Each 4th or 8th grade student will keep a portfolio of his or her work in both writing and mathematics for the year. Criteria have been developed for teachers and students to use in reviewing portfolio work. The criteria are contained in State handbooks, which also contain suggestions for portfolio content. The portfolio approach was designed by a committee that sought input from over 500 Vermont educators. Although the statewide sample will be drawn from grades 4 and 8, all Vermont teachers are being encouraged to implement portfolio assessment.

Writing assessment will be based on a portfolio approach that includes a prompted write sample using a common prompt for all students at a grade level for grades 4, 8 and 11. The writing sample will be holistically scored.

The writing portfolio will contain the following: Table of contents; Best Piece; a dated letter from the student explaining why the choice of the Best Piece and the process of its composition; a poem, play, short story or personal narration; a "personal response" to a cultural or sports event, book, mathematics problem or

BEST COPY AVAILABLE

current issue; and prose pieces from classes other than English and language arts. Writing criteria include: the degree to which the organization suits the writer's purpose, the writing exhibits a sense of personal expression, use of detail adds to clarity, awareness of appropriate usage of, mechanics and grammar, sufficient variety to challenge all students, evidence of teacher-peer response to student drafts, evidence of progress over time and evidence of opportunities for students to revise their work.

The math portfolio will contain a range of materials that will allow students to demonstrate their ability to learn and understand material beyond the "facts and Knowledge" level. Such materials could include : a solution to a problem assigned as homework; a problem made up by the student; a paper done for another subject that includes math, such as analysis of data presented in a graph; or entries from a journal.

The "Best Piece" assessment is based on student selected works from their portfolios that represent their best efforts for the year.

The assessment is conducted by teachers from schools participating in the assessment program. These specially trained teachers evaluate the uniform tests in a central location and visit participating schools to evaluate portfolios and best pieces.

The accountability system was scheduled to be developed state-wide in the 1990-91 school year. Reports were planned for in both writing and mathematics, based on a select sample of schools selected to participate in the pilot year. In 1991-92 it is anticipated that each review team will write reports on each school they visit. Results will be reported back to the schools in early June. Results will be reported from each school on School Report Day.

COMMENT:

This program is in its "pilot" year in 1990-91. It appears to be exemplary in design and could provide a worthwhile model for combining standardized and alternative forms of assessment. Further information on the successes and problems encountered will be valuable to others who would like to develop similar systems.

BEST COPY AVAILABLE

ORGANIZATION: Portland Public Schools

CONTACT PERSONS: Walter Hathaway, Research and Evaluation
Shirley Glick, Curriculum
501 North Dixon Street
Portland, Oregon 97227

PHONE: (503) 249-2000

ABSTRACT: First of all, Portland is working to adapt its districtwide achievement tests to respond to the press for more authentic testing. The Portland Achievement Levels Tests (PALT's) are developed out of item response theory calibrated item banks which are replenished each testing period with newly-tested items, written locally to correspond to new local curriculum goals. The Curriculum goals are revised on a six- to seven-year cycle, prior to the adoption of new instructional materials. As the curriculum goals are revised, so are the designs of the corresponding PALT's. Curriculum and Evaluation teacher committees identify the goals appropriate for objective testing and propose goal-related changes in testing formats, item styles or content. These changes are "tried out" with students and evaluated carefully for equity and quality before use as part of the PALT program. Recent changes which have met standards of both authenticity and psychometric rigor include:

* Student use of calculators with PALT's in Mathematics: Students have calculators available when responding to Geometry, Measurement, Problem-solving, Mathematical Relationships, and Statistics/Graphing/Probability and Applications items. A short section at the end of each test contains items on Computation and Numeration for which calculators are not used.

* Use of published literature in the Reading PALT's: Passages are selected for literary quality, interest to students, and freedom from gender or cultural bias. Each test will have one or more extended passages -- a complete short story, magazine article, or book chapter -- each with several questions that ask about significant features of the text. "Survey items" -- questions asking students about their reading preferences and habits -- appear at the beginning to each test, to address the first currently adopted goal: "Students will read frequently for pleasure and information."

Large scale Direct Writing Assessment (DWA) in the Portland School District began in 1985-86 in response to the demand for a more authentic assessment of students' writing. In addition to a multiple-choice test on language usage skills (PALT), the Direct Writing Assessment was administered districtwide to students in grade 8 (beginning Winter 1988) and grade 4 (Fall 1989), plus Fall and Spring voluntary assessment for all grades. Both of these local DWA programs were suspended in Fall 1990 when the

Oregon Department of Education began a yearly assessment that included a writing assessment at grades 3, 5, 8 and 11. As local resources permit Direct Writing Assessments will be offered at the remaining grades as a local option.

In Direct Writing Assessment in order to make it as realistic and authentic as possible, individual students write a paper on an assigned topic during one class period for three (preferably consecutive) days. On the first day, students plan and write their rough drafts. The second day students edit and revise their rough drafts, using reference materials if they wish. The third and last day is for proofreading and making final copies.

Papers are then shipped to a scoring site where each paper is independently read and scored twice by trained raters according to specific criteria. The analytical rating guide, developed by Portland teachers, identifies and gives scores to each of the six traits of good writing: Ideas & Content, Organization, Voice, Word Choice, Sentence Fluency, and Conventions. The statewide assessments adds a seventh score for the mode of writing: Descriptive, Narrative, Imaginative, Expository, or Persuasive. The rater assigns a value from 1 (weaknesses) to 5 (strengths); therefore, a rating of 3 indicates a balance of strengths and weaknesses. If the ratings differ by more than one point, the paper is read and scored a third time. All ratings and the students' papers are then returned to the classroom teacher for feedback to students.

Curriculum Development and Services of Portland Public Schools has developed a collection of classroom-based assessments in reading, grades 1-8. They provide information about individual students' growth in a range of reading strategies and across a variety of materials. These instruments complement the Portland Achievement Levels Tests, offering data on such District goals as setting purposes for reading, choosing strategies to accomplish reading goals within given text comprehension, relating information and understanding within and between texts, and reading fluently. Depending on extent of use, classroom, building, or cluster-level data is possible. Included in the collection, Portland Public Schools Reading Assessment: Recording Student Progress, are instructions, sample data collection forms, and recommended frequency of use for 1) primary spelling checklist on letter-sound connections, 2) reading strategies checklist, 3) reading habits survey, 4) reading journal prompts, 5) retelling checklist, and 6) comprehension inventory.

Portfolio Assessment in language arts and mathematics is being implemented in one cluster of eight elementary schools. Teachers are collecting student writing samples, administering the spelling checklist, keeping records of students' independent reading, and sampling text comprehension. In addition, teachers and curriculum specialists in a pilot project are developing sample criteria, procedures, and contents for portfolio assessment in language arts, grades 1-12.